PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/31, 1/21, C07K 14/315, 16/12, A61K 39/09, 39/40

A2

(11) International Publication Number: **WO 98/21337**

(43) International Publication Date: 22 May 1998 (22.05.98)

(21) International Application Number: PCT/US97/20586

(22) International Filing Date: 12 November 1997 (12.11.97)

(30) Priority Data:

60/029,444	12 November 1996 (12.11.96)	US
60/038,086	18 February 1997 (18.02.97)	US
60/059,368	19 September 1997 (19.09.97)	US
60/062,473	16 October 1997 (16.10.97)	US

(71) Applicant *(for all designated States except US)*: REGENTS OF THE UNIVERSITY OF MINNESOTA [US/US]; Morrill Hall, 100 Church Street S.E., Minneapolis, MN 55455 (US).

(71)(72) Applicants and Inventors: HOSTETTER, Margaret, K. [US/US]; 1832 Kenwood Parkway, Minneapolis, MN 55405 (US). CHENG, Qi [CN/US]; 11240 48th Avenue North, Plymouth, MN 55442 (US). FINKEL, David, J. [US/US]; Apartment 3, 414 S.E. Erie Street, Minneapolis, MN 55414 (US). SMITH, Beverly [US/US]; 1973 Sheridan Avenue South, Minneapolis, MN 55405 (US).

(74) Agent: MCCORMACK, Myra, H.; Mueting, Raasch, Gebhardt & Schwappach, P.A., P.O. Box 581415, Minneapolis, MN 55458–1415 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

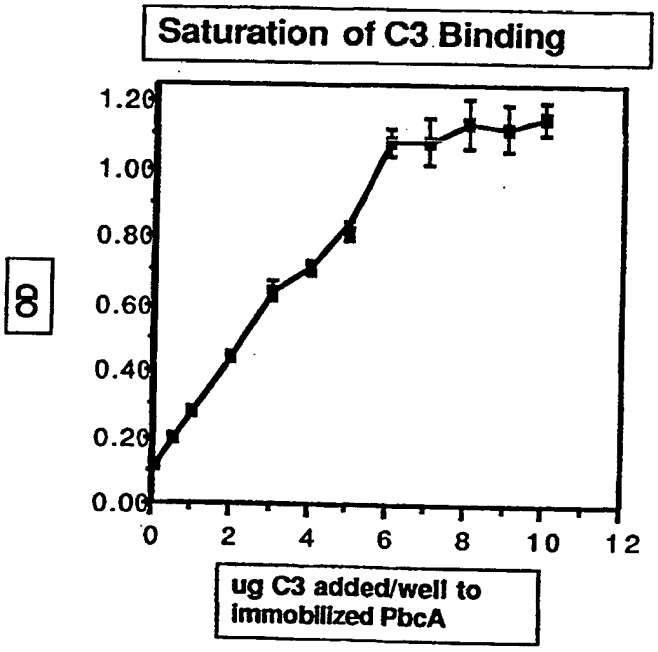
Published
Without international search report and to be republished upon receipt of that report.

(54) Title: C3 BINDING PROTEIN OF *STREPTOCOCCUS PNEUMONIAE*

(57) Abstract

This invention relates to the identification of a human complement C3 binding protein from *Streptococcus pneumoniae* and to its sequence and to methods for its purification and use. The protein binds but does not degrade or cleave C3 and is implicated in *S. pneumoniae* virulence. The protein is recognized by antibodies produced by humans recovering from pneumococcal infection.

PbcA Binds C3 in ELISA



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

C3 BINDING PROTEIN OF *STREPTOCOCCUS PNEUMONIAE*

Field of the Invention

This invention relates to *Streptococcus pneumoniae* and in particular this invention relates to the identification of an *S. pneumoniae* protein that is implicated in *S. pneumoniae* virulence and is capable of binding the complement protein, C3.

Background of the Invention

This application is a continuation in part of a provisional application filed October 16, 1997 (U.S. Serial No. 60/062,473) entitled "C3 BINDING PROTEIN OF *S. PNEUMONIAE*." This application is also a continuation in part of a provisional application filed on September 19, 1997 (U.S. Serial No. 60/059,368), a provisional application filed on February 18, 1997 (U.S. Serial No. 60/038,086) and a provisional application filed on November 12, 1996 (U.S. Serial No. 60/029,444) all entitled "C3 BINDING PROTEIN OF *S. PNEUMONIAE*."

Respiratory infection with the bacterium *Streptococcus pneumoniae* (*S. pneumoniae*) leads to an estimated 500,000 cases of pneumonia and 47,000 deaths annually. Those persons at highest risk of bacteremic pneumococcal infection are infants under two years of age and the elderly. In these populations, *S. pneumoniae* is the leading cause of bacterial pneumonia and meningitis. Moreover, *S. pneumoniae* is the major bacterial cause of ear infections in children of all ages. Both children and the elderly share defects in the synthesis of protective antibodies to pneumococcal capsular polysaccharide after either bacterial colonization, local or systemic infection, or vaccination with purified polysaccharides. *S. pneumoniae* is the leading cause of invasive bacterial respiratory disease in both adults and children with HIV infection and produces hematogenous infection in these patients (Connor et al. *Current Topics in AIDS* 1987;1:185-209 and Janoff et al. *Ann. Intern. Med.* 1992;117(4):314-324).

Individuals who demonstrate the greatest risk for severe infection are not able to make antibodies to the current capsular polysaccharide vaccines. As a result, there are now four conjugate vaccines in clinical trial. Conjugate vaccines consist of pneumococcal capsular polysaccharides coupled to protein carriers or adjuvants in an attempt to boost the antibody response. However, there are other potential problems with conjugate vaccines currently in clinical trials. For example, pneumococcal serotypes that are most prevalent in the United States are different from the serotypes that are most common in places such as Israel, Western Europe, or Scandinavia. Therefore, vaccines that may be useful in one geographic locale may not be useful in another. The potential need to modify currently available capsular polysaccharide vaccines or to develop protein conjugates for capsular vaccines to suit geographic serotype variability entails prohibitive financial and technical complications. Thus, the search for immunogenic, surface-exposed proteins that are conserved worldwide among a variety of virulent serotypes is of prime importance to the prevention of pneumococcal infection and to the formulation of broadly protective pneumococcal vaccines. Moreover, the emergence of penicillin and cephalosporin-resistant pneumococci on a worldwide basis makes the need for effective vaccines even more exigent (Baquero et al. *J. Antimicrob. Chemother.* 1991;28S;31-8).

Several pneumococcal proteins have been proposed for conjugation to pneumococcal capsular polysaccharide or as single immunogens to stimulate immunity against *S. pneumoniae*. Surface proteins that are reported to be involved in adhesion of *S. pneumoniae* to epithelial cells of the respiratory tract include PsaA, PspC/CBP112, and IgA1 proteinase (Sampson et al. *Infect. Immun.* 1994;62:319-324, Sheffield et al. *Microb. Pathogen.* 1992; 13: 261-9, and Wani, et al. *Infect. Immun.* 1996; 64:3967-3974). Antibodies to these adhesins could inhibit binding of pneumococci to respiratory epithelial cells and thereby reduce colonization. Other cytosolic pneumococcal proteins such as pneumolysin, autolysin, neuraminidase, or hyaluronidase are proposed as vaccine antigens because antibodies could potentially block the toxic effects of

these proteins in patients infected with *S. pneumoniae*. However, these proteins are typically not located on the surface of *S. pneumoniae*, rather they are secreted or released from the bacterium as the cells lyse and die (Lee et al. *Vaccine* 1994; 12:875-8 and Berry et al. *Infect. Immun.* 1994; 62:1101-1108). While use of these cytosolic proteins as immunogens might ameliorate late consequences of *S. pneumoniae* infection, antibodies to these proteins would neither promote pneumococcal death nor prevent pneumococcal colonization.

A prototypic surface protein that is being tested as a pneumococcal vaccine is the pneumococcal surface protein A (PspA). PspA is a heterogeneous protein of about 70-140 kDa. The PspA structure includes an alpha helix at the amino terminus, a proline-rich sequence in the mid-portion of the protein, and terminates in a series of choline-binding repeats at the carboxy-terminus. Although much information regarding its structure is available, PspA is not structurally conserved among a variety of pneumococcal serotypes, and its function is entirely unknown (Yother et al. *J. Bacteriol.* 1992;174:601-9 and Yother *J. Bacteriol.* 1994;176:2976-2985). Studies have confirmed the immunogenicity of PspA in animals (McDaniel et al. *Microb. Pathogen.* 1994; 17;323-337). Despite the immunogenicity of PspA, the heterogeneity of PspA, its existence in four structural groups (or clades), and its uncharacterized function complicate its ability to be used as a vaccine antigen.

In patients who cannot make protective antibodies to the type-specific polysaccharide capsule, the third component of complement, C3, and the associated proteins of the alternative complement pathway constitute the first line of host defense against *S. pneumoniae* infection. Because complement proteins cannot penetrate the rigid cell wall of *S. pneumoniae*, deposition of opsonic C3b on the pneumococcal surface is the principal mediator of pneumococcal clearance. Interactions of pneumococci with plasma C3 are known to occur during pneumococcal bacteremia, when the covalent binding of C3b, the opsonically active fragment of C3, initiates phagocytic recognition and ingestion (Johnston et al. *J. Exp. Med* 1969;129:1275-1290, Hasin HE, *J. Immunol.* 1972; 109:26-31 and Hostetter et al. *J. Infect. Dis.* 1984; 150:653-61).

C3b deposits on the pneumococcal capsule, as well as on the cell wall. This method for controlling *S. pneumoniae* infection is fairly inefficient and could be beneficially amplified by the presence of antibodies to surface components of *S. pneumoniae*. There currently exists a strong need for methods and therapies to limit *S. pneumoniae* infection.

Summary of the Invention

The present invention relates to the identification and purification of an about 90 kDa to about 110 kDa (± 5 kDa) protein, as determined following electrophoresis on a 15% SDS-PAGE gel. The protein is named PbcA and is isolatable from *S. pneumoniae* strains that are capable of binding to human complement protein C3. The protein, PbcA, comprises an amino terminus containing region comprising SEQ ID NO:1 and is capable of binding but not cleaving or degrading the human complement protein C3. The protein also comprises a proline rich region and in one embodiment is a surface exposed protein of *S. pneumoniae*.

This invention also relates to the production of antibodies specifically recognizing PbcA. In one embodiment the antibodies are polyclonal and in another embodiment the antibodies are monoclonal. The antibodies can be produced by immunizing a mammal with all or a portion of PbcA. In one embodiment, the monoclonal antibodies are rodent derived.

In another aspect of this invention a method is provided for generating an immune response to *S. pneumoniae in vivo* comprising the steps of: administering a protein or an immunogenic fragment of a protein from *S. pneumoniae* to an animal wherein the amino terminus containing region of the protein comprises SEQ ID NO:1. In one embodiment, the protein is capable of binding but not cleaving or degrading the human complement protein C3. Preferably, the method further comprises detecting an immune response to *S. pneumoniae* in the mammal. Preferably, the immune response comprises the production of antibodies to *S. pneumoniae*. The animal can be a mouse, rat, chinchilla, a rabbit or a human and the method can further comprise the steps of

isolating antibody producing cells from the mammal and preparing monoclonal antibodies to the C3 binding protein.

In yet another aspect of the invention a method is disclosed for obtaining a purified C3 binding protein from *S. pneumoniae* comprising the steps of: obtaining a protein sample from *S. pneumoniae*; precipitating the protein to form a precipitate; applying the precipitate to a Thiopropyl Sepharose 6B affinity chromatography column comprising methylamine-treated C3; and eluting the C3 binding protein from the column using an elution buffer comprising about 20% ethanol. In one aspect of this embodiment the invention relates to C3 binding protein preparable by these methods.

This invention also relates to a C3 binding protein having the sequence of SEQ ID NO:6 and to isolated nucleic acid encoding a C3 binding protein and comprising the DNA sequence of SEQ ID NO:5 and to isolated nucleic acid having the DNA sequence of SEQ ID NO:4.

In another aspect of this invention, the invention relates to isolated nucleic acid encoding C3 binding protein of about 90 kDa to about 110 kDa (± 5 kDa), in one embodiment, and comprising nucleic acids 1-1500 of SEQ ID NO:5 and to an isolated nucleic acid sequence encoding the C3 binding protein wherein the protein exhibits C3 binding activity and wherein the protein comprises at least 80% nucleic acid homology to nucleic acids 1-1500 of SEQ ID NO:5. Preferably, the nucleic acid homology is at least 95%.

The invention also relates to a C3 binding protein isolatable from *S. pneumoniae* having DNA that is hybridizable to a nucleic acid fragment of at least 500 bp from nucleic acids 1-1500 of SEQ ID NO: 5 under hybridization conditions of about 6X SSC, 5X Denhardt's, 0.5% SDS, 100 µg/ml denatured, fragment salmon sperm DNA overnight at 65°C and washed in 2X SSC, 0.1%SDS, one time at room temperature for about 10 mn, followed by one time at 65°C for about 15 mn and followed by at least one wash in 0.2 X SSC, 0.1% SDS at room temperature for about 3-5 minutes. Preferably this protein further comprises at least 2 choline binding repeat and still more preferably the protein further comprises at least 2 choline binding repeats.

The invention also relates to peptide fragments of at least 15 bp from SEQ ID NO:5 and to insertion and deletion mutants that do not express PbcA.

In another aspect of this invention, the invention relates to PbcA proteins. In one embodiment the invention relates to an isolated protein comprising SEQ ID NO:1 and at least two choline binding repeats. Preferably the protein is isolated from *S. pneumoniae* and also preferably the protein binds human complement protein C3. In one version of this embodiment, the protein is a recombinant protein or a purified protein from *S. pneumoniae*. Preferably the protein has a molecular weight as determined on a 15% polyacrylamide gel of between about 90 kDa to about 110 kDa and preferably the protein comprises a proline rich region. Additionally the protein can comprise SEQ ID NO:2 or SEQ ID NO:3.

Alternatively, the protein of this invention can comprise SEQ ID NO:6 or the protein can be an isolated protein capable of binding to, but not cleaving or degrading, human complement C3 and wherein the protein comprises SEQ ID NO:1. Preferably the protein is isolated from *S. pneumoniae* and in one embodiment, the protein further comprises a proline rich region. The protein can further comprise SEQ ID NO:2. Preferably the protein has at least about 95% homology to a C3 binding protein from *S. pneumoniae* and also preferably, the protein has a molecular weight as determined on a 15% polyacrylamide gel of between about 90 kDa to about 110 kDa. In one aspect, the isolated protein is a recombinant protein and in another, the protein is isolated from *S. pneumoniae* bacteria.

In another embodiment of the proteins of this invention, the invention relates to a recombinant protein comprising SEQ ID NO:1, wherein the protein has a molecular weight as determined on a 15% polyacrylamide gel of between about 90 kDa to about 110 kDa. Preferably the protein binds human complement protein C3. The protein can further include a proline rich region and preferably the protein does not cleave or degrade human complement protein C3.

In another embodiment of the proteins of this invention, the protein comprises amino acids 1-410 of SEQ ID NO:6.

In yet another embodiment of the proteins of this invention, the invention relates to a protein that binds, but does not cleave or degrade, human complement protein C3, wherein nucleic acid encoding the protein hybridizes to SEQ ID NO:4 under hybridization conditions of 6X SSC, 5X Denhardt, 0.5% SDS, and 100 μg/ml fragmented and denatured salmon sperm DNA hybridized overnight at 65°C and washed in 2X SSC, 0.1% SDS one time at room temperature for about 10 minutes followed by one time at, 65°C for about 15 minutes followed by at least one wash in 0.2XSSC, 0.1% SDS at room temperature for at least 3-5 minutes. Preferably, the protein further comprises SEQ ID NO:1 and optionally, the protein can comprise SEQ ID NO:2 or SEQ ID NO:3. Preferably the protein is at least 15 amino acids in length. Preferably the protein comprises a proline rich region. In one aspect of this embodiment, the protein is a recombinant protein and in another, the protein is a synthetic peptide. In one aspect of this embodiment, the protein is a peptide of at least 15 amino acids from SEQ ID NO:6. The proteins can also be used to create antibody and the proteins of this embodiment can be used to generate antibody capable of specifically binding to the protein. In one embodiment, the antibody is a monoclonal antibody and in another the antibody is a polyclonal antibody. Preferably the monoclonal antibody is at least partially rodent-derived.

This invention also relates to nucleic acid encoding the proteins of this invention. In one embodiment, the nucleic acid of this invention encodes a protein comprising at least two choline binding domains and SEQ ID NO: 1. Preferably the protein encoded by the nucleic acid further comprising a proline rich region. Also preferably, the nucleic acid is isolated from an *S. pneumoniae* genome. Preferably, the nucleic acid is capable of hybridizing to SEQ ID NO:4 and in another embodiment, the protein encoded by the nucleic acid binds to human complement protein C3. In one aspect of this embodiment, the nucleic acid is positioned in a nucleic acid vector. Preferably the vector is an expression

vector and the expression vector directs expression of the protein by the nucleic acid.

In another embodiment of the nucleic acid of this invention, the invention relates to isolated nucleic acid encoding a protein comprising SEQ ID NO:1 and a proline rich region wherein the protein encoded by the nucleic acid binds but does not cleave or degrade human complement C3.

In yet another embodiment, the invention relates to isolated nucleic acid fragment encoding an about 90 kDa to about 110 kDa protein with C3 binding activity, wherein the nucleic acid fragment has at least 80% homology to at least 500 bp from nucleic acids 1-1500 of SEQ ID NO:5 and in another embodiment, the isolated nucleic acid fragment comprises base pairs 1-1500 of SEQ ID NO:5.

The invention also relates to a method for isolating a C3 binding protein from a bacterium comprising the steps of: obtaining a protein sample from a bacterium; applying the sample to a solid support comprising methylamine treated complement protein C3; washing the solid support; and removing a C3 binding protein from the solid support in a solution comprising alcohol wherein the C3 binding protein does not cleave or degrade C3. Preferably the bacterium is *S. pneumoniae* and in another embodiment, the bacterium is *E. coli*. Preferably the solid support comprises an affinity column and preferably the alcohol is ethanol. In one embodiment, the solution comprising alcohol is a buffer comprising 20% ethanol. The invention also relates to C3 binding protein preparable by this method.

The invention also relates to a method for producing an immune response to *S. pneumoniae* comprising the steps of: administering a therapeutically effective amount of at least a portion of a protein to a mammal, wherein the protein binds but does not cleave or degrade human complement protein C3 and, wherein nucleic acid encoding the protein hybridizes to SEQ ID NO:4 under hybridization conditions of 6XSSC, 5X Denhardt, 0.5% SDS, and 100 µg/ml fragmented and denatured salmon sperm DNA, hybridized overnight at 65°C and washed in 2x SSC, 0.1% SDS one time at room temperature for

about 10 minutes followed by one time at 65°C for about 15 minutes followed by at least one wash in 0.2xSSC, 0.1% SDS at room temperature for at least 3-5 minutes; and detecting an immune response to the protein. In one embodiment, the protein is at least 15 amino acids in length and in another the protein is a chimeric protein. In yet another embodiment, the protein comprises SEQ ID NO:1. In another embodiment, the protein has a molecular weight on a 15% polyacrylamide gel of between about 90 kDa to about 110 kDa.

The invention also relates to a method for reducing *S. pneumoniae* binding to C3 comprising the steps of: administering a therapeutically effective amount of at least a portion of an antibody to a mammal, wherein the antibody specifically recognizes a C3 binding protein from *S. pneumoniae*, and wherein the C3 binding protein is a protein that binds, but does not cleave or degrade human complement protein C3. In one embodiment the antibody comprises at least one variable domain from a monoclonal antibody. In another embodiment, the antibody is administered to the air passages of the mammal or intravenously.

In yet another aspect of this invention, the invention relates to a non-naturally occurring *S. pneumoniae* bacterium that does not express a detectable human complement C3 binding protein, wherein the C3 binding protein comprises SEQ ID NO:1. In one embodiment, the bacterium is the product of an insertion into the gene encoding the C3 binding protein and in another, the bacterium is a product of a deletion in the gene encoding the C3 binding protein. In yet another embodiment, the bacterium is produced by homologous recombination of the gene encoding the C3 binding protein gene with at least a portion of a non-native C3 binding protein gene and in one aspect of this embodiment, the gene encoding at least a portion of the non-native C3 binding protein comprises a mutation within the non-native gene.

Brief Description of the Figures

Fig. 1 is an amino acid alignment of the choline binding regions of PbcA from *S. pneumoniae* strains CP1200, R6x and 23F.

Fig. 2 is a graph illustrating PbcA binding to human C3 by ELISA.

Fig. 3 summarizes studies to assess the ability of antibody to PbcA to inhibit C3 binding. Figure 3a is a Western blot illustrating the specificity of the antibody preparation for PbcA from CP1200. Figure 3b is a graph illustrating the ability of antibody to PbcA to inhibit C3 binding by ELISA.

Fig. 4 illustrates an exemplary strategy for producing a *pbc*A insertion mutant.

Fig. 5a is a photograph of a Southern blot of DNA from transformed and nontransformed *S. pneumoniae* cells probed with a 1.5kb fragment of *pbc*A. Figure 5b is a photograph of a Southern blot probed with vector pVA891.

Figure 6 is a photograph of Western blot experiments demonstrating that the insertion-duplication mutants of this invention do not produce PbcA protein capable of binding C3 or antibody to PbcA.

Detailed Description of the Preferred Embodiments

The present invention provides a *S. pneumoniae* protein that binds C3. The protein is about 90 kDa to about 110 kDa (± about 5 kDa) in mass as observed following electrophoresis on a 15% SDS-PAGE gel and is referred to in this disclosure as "PbcA." PbcA was initially identified using the methods of Example 1. Supernatant proteins from exponentially growing cultures of *S. pneumoniae* were tested by Western blot for their ability to bind to purified human C3 following electrophoresis on a 15% SDS-PAGE gel. In the *S. pneumoniae* strain CP1200, both lysates and supernatants from exponentially growing cells produced a protein band, identified as PbcA, at approximately 90 kDa on 15% SDS-PAGE gels under non-reducing conditions. The protein bound methylamine-treated human C3 (produced as described by Hostetter MK, et al. *J. Inf. Dis.* 150:653-661, 1984), labeled with biotin. PbcA is present in lysates and supernatants from a variety of other *S. pneumoniae* strains. The size of PbcA

varies on a 15% SDS-PAge gel in *S. pneumoniae* strains from about 90 kDa to about 110 kDa. All PbcA proteins identified thus far are capable of binding, but not cleaving or degrading, purified human complement protein C3. In some strains protein bands of smaller molecular weight are observed.

As will be recognized by those of skill in the art, there are a variety of methods for isolating an individual protein from bacterial supernatants. As one example, the *S. pneumoniae* C3 binding protein, PbcA, is isolated using the methods of Example 2. PbcA, can be purified from other secreted pneumococcal proteins by affinity chromatography. Secreted proteins can be precipitated in a final concentration of 10% trichloroacetic acid (TCA) at 4°C overnight according to Example 2. Resuspended proteins from the TCA precipitate are subjected to affinity column chromatography using methylamine-treated human C3 (*supra*). Elution of the PbcA protein from the affinity column has proven to be difficult. Surprisingly, the PbcA protein from the column can be eluted using an elution buffer comprising an alcohol, preferably ethanol and more preferably about 20% ethanol in the Tris-HCl/NaCl wash buffer. In view of this disclosure, now that PbcA has been identified, those skilled in the art will recognize that other methods could be used to identify, isolate and purify the protein from a variety of C3-binding *S. pneumoniae* without undue experimentation.

Multiple eluates can be pooled to obtain sufficient sample for further analysis. As one example, a sample can be electrophoresed on an SDS-PAGE gel and transferred to nitrocellulose. The protein can be subjected to amino terminal analysis and tryptic digestion for internal peptide sequencing. The following sequences were obtained from the tryptic digest analysis:

A peptide positioned near the amino terminus:

TENEGSTQAATSSNMAKTEH (SEQ ID NO:1)

And internal regions:

EKPAEQPQPAPATQP (SEQ ID NO:2)

SSDSSVGEETLPSSSLK (SEQ ID NO:3)

SEQ ID NO:2 is proline rich and has at least a 75% homology with the *S. pneumoniae* protein PspA over 13 amino acids. Although the proline rich region of PspA aligned with SEQ ID NO:2, neither SEQ ID NO:1 nor SEQ ID NO:3 had any substantial homology to any proteins or peptides previously published in the GenBank database (less than 35% homology). The term "proline rich" as used herein refers to a protein having a stretch of amino acids having at least 5 proline amino acids over a total of about 15 amino acids.

These sequences were confirmed in *S. pneumoniae* strain CP1200 following isolation of the gene and sequencing to obtain the nucleic acid sequence encoding PbcA. In one embodiment of this invention, PbcA is an about 90 kDa to about 110 kDa (± 5 kDa, meaning about 85 kDa to about 115 kDa) when *S. pneumoniae* proteins are separated on a 15% SDS-PAGE gel, and in another embodiment, the protein further includes SEQ ID NO:2 and SEQ ID NO:3.

Oligonucleotides corresponding in whole or in part to SEQ ID NOS: 1-3 are useful for identifying and isolating the nucleic acid encoding PbcA (the gene encoding PbcA is termed *pbcA*) and for isolating the *pbc*A and PbcA from a variety of *S. pneumoniae* strains. For example, oligonucleotides corresponding in whole or in part from SEQ ID NOS 1-3 can be used to amplify sequences from genomic DNA isolated from *S. pneumoniae* using standard polymerase chain reaction technology. The amplified sequences can then be directly used as probes or the amplified sequences can be incorporated into a vector for plasmid amplification in a suitable host such as a bacteria or virus and then isolated for sequencing, cloning and for use as probes to detect DNA from libraries of *S. pneumoniae*. The DNA isolated from these procedures is useful in sequencing reactions to obtain the nucleic acid sequence encoding PbcA and to produce vectors, such as expression vectors encoding PbcA as well as for producing recombinant protein. Example 5 provides a preferred method for isolating nucleic acid encoding PbcA.

PbcA can be expressed as a recombinant protein or isolated from S. pneumoniae lysates. The *S. pneumoniae* C3 binding protein, PbcA, binds C3

without cleaving or degrading the C3 molecule. A number of bacterial proteins have been reported to bind and to cleave C3 or other complement proteins. For example, a 140 kDa C5a peptidase from group A streptococci cleaves a His_{67}-Lys_{68} bond at the carboxy terminus of C5a, thereby abolishing the chemoattractant capabilities of the molecule. An enzyme related to the C5a peptidase is also found in group B streptococci (Cleary et al. *Infect. Immun.* 1992; 60:4239-4244 and Bohnsack et al. *Biochim. et Biophys. Acta* 1991; 1079:222-228). Production of an elastase-like enzyme, as can be seen with 24-hour culture supernatants from *Pseudomonas aeruginosa* (Suter et al. *J. Infect. Dis.* 1984; 149:523-31), cleaves C3 into characteristic fragments of 66 kDa and 100 kDa from the C3 α-chain. Like the elastase-like enzyme from *P. aeruginosa*, a 56 kDa neutral cysteine proteinase from *Entamoeba histolytica* cleaves C3a between residues Ser_{78}/Asn_{79}, yielding a defined C3 cleavage fragment of 105 kDa (Reed, et al. *J. Immunol.* 1989; 143:189-95).

C3-cleaving proteinases have been isolated from the membranes of some mammalian cells, including human erythrocytes (p57), neutrophils, and melanoma cells resistant to complement-mediated killing. These proteins are typically serine proteases which yield defined cleavage fragments. For example, p57 cleaves both the α- and β-chains of C3, while the melanoma proteinase cleaves only the α'-chain of C3b, generating a fragment of 35 kDa (Charriaut-Marlangue et al. *Biochem. Biophys. Res. Commun.* 1986; 140:1113-1120 and Ollert et al. *J. Immunol.* 1990; 144:3862-7).

In contrast to these studies, C3 cleavage and/or degradation was not observed with PbcA. PbcA binds C3 without the production of defined cleavage fragments and without evidence of degradation of C3. Although there are microbial precedents for binding and cleaving of complement proteins, there is no previously reported microbial protein that binds C3 non-covalently without degrading or cleaving the molecule. The interaction of C3 and other complement components with proteins from group A and B streptococci, *P. aeruginosa*, and amoebae appears to be quite distinct from what has been observed with *S. pneumoniae*.

Table 1 (SEQ ID NO:4) provides a nucleic acid sequence encompassing the open reading frame encoding a 90 kDa PbcA protein from *S. pneumoniae* strain CP1200. Untranslated 5' and 3' regions are also included in SEQ ID NO:4. The open reading frame encoding PbcA begins at nucleotide 383 and ends with nucleotide 2074. Table 2 is a map providing the amino acids (SEQ ID NO: 6) encoded by the nucleic acids of the major open reading frame from SEQ ID NO:4 (provided in Table 2 as SEQ ID NO:5). The protein predicted from SEQ ID NO:4 contains an amino terminus containing segment with C3 binding activity (upstream from the choline binding repeat region) and a series of choline binding repeats (beginning in SEQ ID NO:5 at about nucleic acid position 1501).

Primers were selected from SEQ ID NO:4 to span the choline binding repeats to assess choline binding repeat variability between strains. The primers used were:

5' GCACAACCATCTACTCCA 3' (SEQ ID NO: 7), and

5' GTACAGGAATTCAGTATTAACTA 3' (SEQ ID NO:8)

Amplification reactions were performed using DNA from three different *S. pneumoniae* strains: CP1200, R6x and virulent strain 23F (obtained from Dr. Steve Pelton, Boston City Hospital, Boston, MA and identified as isolate "freezer #365"). The results of the amplification studies indicated that the number of choline binding repeats varied depending on the *S. pneumoniae* strain. For example, strain CP1200 contained about 4 repeats while 23F contained about 8 repeats and R6X contained at least about 10 choline binding repeats using the Yother et al. model for choline binding repeat regions (see *infra*). All *S. pneumoniae* strains studied thus far have at least two choline binding repeats. Therefore, in another embodiment of this invention, PbcA is a C3-binding protein from *S. pneumoniae* including SEQ ID NO:1 and at least 2 choline binding repeats.

Choline binding domains are known in the art and a number of references discuss the characteristics of a variety choline binding domains (see, for example, Du Clos et al. *J. Biol. Chem.* 266(4):2167-2171, 1991; Liu et al. *J.*

Biol. Chem. 266(22):14813-14821, 1991; Agrawal et al. *J. Biol.Chem.* 267(35):25352-25358, 1992; and *Nature Structural Biology* 3(4):346-354, 1994). A choline-binding repeat sequence has been identified in *S. pneumoniae* protein PspA as TGWKQENGMWYFYNTDGSMA (see Yother, J. and White, JM, *J. Bacteriology* 176:2976-2985, 1994) Choline binding repeats are associated with membrane binding proteins. Without intending to limit this invention, the strains studied thus far indicate that the virulent strains appear to have more choline binding repeats on average than nonvirulent strains. At the very least it appears that *S. pneumoniae* strains show considerable variability in the choline binding repeat region.

These results are consistent with electrophoretic studies assessing variability in the size of PbcA proteins obtained from a number of *S. pneumoniae* strains. Proteins isolated according to the methods of this invention, when separated by SDS-PAGE, demonstrate some size variability (See Fig. 6) from about 90 kDa to about 110 kDa (± 5 kDa). This variability can be attributed, at least in part, to the variability in length of the choline binding repeat region. The proteins of this invention also include a series of peptides at the carboxy terminus of the choline binding repeat region.

Fig. 1 provides a comparison of the amino acid sequence from a region containing the choline binding repeat region from PbcA for three different *S. pneumoniae* strains: CP1200, R6x and 23F(BD23). The amino acids in the boxes represent variations in the 23F (virulent strain BD23) choline binding domain repeat region.

One example of a fragment containing nucleic acid encoding PbcA protein is provided in SEQ ID NO:4. The nucleic acid sequence encoding PbcA in strain CPI200 is provided as SEQ ID NO:5 (see Table 2). In another aspect of this invention, a protein of this invention has the amino acid sequence of SEQ ID NO:6. In one embodiment, a protein of this invention includes amino acids 1-410 of SEQ ID NO:6. In addition, variability between strains has been identified at the amino acid and nucleic acid level. For example, the amino terminus contains some variability and SEQ ID NO:1 may lack the first

threonine residue. In general, PbcA proteins include C3 binding proteins from *S. pneumoniae* that preferably have at least 80% nucleic acid homology within the DNA of the amino terminus-containing region (the region amino-terminal to the choline binding repeats) to SEQ ID NO:5. More preferably, the PbcA proteins have at least 95% homology to the amino terminus-containing region of PbcA and still more preferably the PbcA protein includes SEQ ID NO:1. In another embodiment the PbcA proteins additionally include SEQ ID NO:2 or SEQ ID NO:3.

In one example, the nucleic acid encoding PbcA can be obtained from a variety of *S. pneumoniae* strains. A *S. pneumoniae* genomic library can be prepared using *S. pneumoniae* genomic DNA and in a preferred example, an *S. pneumoniae* genomic library was prepared using the CP1200 strain. Custom libraries can be obtained using a variety of standard methods for library construction. In these studies genomic DNA from *S. pneumoniae* was given to a commercial custom library supplier (Stratagene, LaJolla, CA). In one example, *S. pneumoniae* CP1200 strain genomic DNA was used to prepare the library (Example 5) The results of these studies identified a cloned nucleic acid fragment encoding PbcA. This fragment is useful for identifying PbcA encoding nucleic acid in other *S. pneumoniae* strains and the nucleic acid can be incorporated into vectors including expression vectors, for example to produce recombinant protein using methods such as those described by Sambrook et al. (cited below).

PbcA is preferably encoded by nucleic acid that is capable of hybridizing to at least 500 bp from the amino terminus region of SEQ ID NO: 5 under hybridization conditions of about 6X SSC, 5X Denhardt's, 0.5% SDS, 100 µg/ml denatured, fragmented salmon sperm DNA hybridized overnight at 65°C and washed in 2X SSC, 0.1%SDS, one time at room temperature for about 10 mn, followed by one time at 65°C for about 15 mn and followed by at least one wash in 0.2 X SSC, 0.1% SDS at room temperature for about 3-5 minutes. The protein of this invention preferably includes at least two and preferably at least three choline binding repeats.

This invention also relates to nucleic acid fragments of at least 20 bp from SEQ ID NO:5 and to isolated nucleic acid fragments of at least 100 base pairs that hybridize to the *S. pneumoniae* genome or to SEQ ID NO:5 under the hybridization conditions of about 6X SSC, 5X Denhardt's, 0.5% SDS, 100 µg/ml denatured, fragment salmon sperm DNA hybridized overnight at 65°C and washed in 2X SSC, 0.1%SDS, one time at room temperature for about 10 min., followed by one wash at 65°C for about 15 min and followed by at least one wash in 0.2 X SSC, 0.1% SDS at room temperature for about 3-5 minutes. Preferably the nucleic acid fragments of at least 100 base pairs are part of a nucleic acid sequence that encodes a C3 binding protein that is at about 90 kDa to about 110 kDa. Preferably the nucleic acid fragment encodes a protein having at least two choline binding repeats.

As demonstrated in Example 3, the immunogenicity of PbcA can be studied by testing for the presence of antibodies to PbcA in acute and convalescent sera from patients with culture-proven pneumococcal infection. *S. pneumoniae* proteins can be separated by SDS-PAGE electrophoresis, transferred to nitrocellulose, and incubated with acute or convalescent serum (standard Western blot procedure). Results to detect the presence of antibodies to PbcA in human sera indicated that acute serum did not contain antibodies to any pneumococcal proteins released into the supernatant but that convalescent serum contained antibodies that recognized a band on an acrylamide gel with a molecular weight of PbcA.

PbcA, or fragments of PbcA, can be used to produce antibodies specific to PbcA. The term "specific" is used to mean that when the antibodies to PbcA or fragments thereof are used in Western blots containing PbcA or a PbcA fragment or peptide under standard Western Blot assay conditions, the antibodies recognize only PbcA or its degradation or truncated products. Purified PbcA or fragments, preferably of at least about 15 amino acids in length, can be used to inject laboratory animals for the production of polyclonal and monoclonal antibodies to PbcA. Those skilled in the art will recognize that the methods for producing polyclonal antibodies and for producing monoclonal

antibodies are known in the art and include the methods disclosed by Harlow et al. (cited below). Purified antibodies can be generated using the isolated protein of this invention, or fragments thereof, without undue experimentation. The antibodies are useful in *in vitro* assays to test for the presence or absence of PbcA protein and to test for the ability of the antibodies to block C3 binding as well as in *in vivo* assays to test for the ability of the antibodies to provide passive protection against pneumococcal infection, whether local or systemic. Antibody fragments and chimeric antibodies can be used and the antibodies include at least one variable domain from an antibody specifically recognizing PbcA.

Moreover, purified proteins, peptides and polypeptides from PbcA can be injected into animals, and later humans, to produce an antibody response to *S. pneumoniae*. As used herein, the terms proteins, peptides and polypeptides are used interchangeably. Therefore, for purposes of this application and as used in the claims, a protein refers to proteins, protein fragments, peptides and polypeptides. Methods for introducing the protein, peptide or polypeptide fragment of PbcA to a mammal with an appropriate adjuvant, if necessary, are known in the art. Moreover, all or a portion of PbcA can be produced as an isolated protein or as a recombinant protein. Recombinant proteins can include all or a part of PcbA or can be formed as a chimeric protein. As used herein, the term "chimeric protein" refers to a recombinant protein including all or at least 15 amino acids of PcbA and amino acid sequence from at least one other protein positioned amino to, carboxy to, or on either side of the PbcA-derived amino acid. The 15 amino acids of PbcA of the chimeric protein are preferably unique to *S. pneumoniae*-derived proteins. Since patients with cleared *S. pneumoniae* infection have convalescent antibodies recognizing PbcA, it is known that the immune system can mount an immune response that includes the production of antibodies to PbcA. Ultimately, PbcA can be used as an immunogen for a pneumococcal vaccine. Further, based on the findings of these studies, antibody produced from the isolated protein, PbcA, can be used in Western Blot analyses to determine whether or not a particular *S. pneumoniae*

strain's virulence or avirulence has been associated with the presence or absence of PbcA.

As demonstrated in Example 4, PbcA is implicated in pneumococcal virulence (see Example 4). In this example an avirulent 23F pneumococcal isolate was inoculated into the ears of chinchillas. Virulent pneumococcal isolates typically cause otitis media and the influx of leukocytes after injection at concentrations as low as 1×10^2 colony forming units (cfu). In contrast, the avirulent 23F strain was inoculated at concentrations of up to about 1×10^7 cells and even at that level was unable to cause otitis media or inflammation (Giebink et al. *J. Infect. Dis.* 1993; 167:347-355). As discussed in Example 4, studies disclosed here indicate that the avirulent strain reported by Giebink lacked detectable PbcA protein by Western blot. To further study the implications of PbcA on virulence, an insertion/duplication mutation of PbcA was prepared for further study (see Example 8).

The present invention provides a detailed method of purification of PbcA and studies indicat that PbcA is: (a) immunogenic in man; (b) conserved within the mass range of about 90 kDa to about 110 kDa (as observed on a 15% SDS-PAGE gel) among a variety of pneumococcal serotypes (that is, PbcA has been identified in a variety of serotypes); and (c) absent in an avirulent 23F strain that is incapable of causing otitis media in a chinchilla model.

To demonstrate that PbcA can be detected in an ELISA assay, PbcA coated wells of an ELISA plate were incubated with antibody prepared to purified PbcA from *S. pneumoniae* strain CP1200.

To demonstrate the ability of PbcA to bind human C3, purified PbcA from CP1200 was used to coat ELISA plates and the coated protein was incubated with methylamine-treated human C3. The binding of methylamine-treated human C3 was assessed using antibody to human C3 conjugated with horeseradish peroxidase (See Example 7). Figure 2 is a graph illustrating the ability of PbcA to bind increasing concentrations of human C3. Specific antibody to PbcA blocked the binding of purified human C3 to PbcA in a dose-dependent fashion. These experiments are detailed in Example 7 and in Figure

3. The Western blot (Figure 3a) demonstrates that the IgG fraction of antibody to PbcA blocked the binding of human C3 to PbcA immobilized on nitrocellulose. The graph (Figure 3b) demonstrates that affinity-purified antibody to PbcA blocked the binding of human C3 to PbcA immobilized on ELISA plates.

A PbcA insertion construct was prepared to interrupt a *pbc*A gene in CP1200 and R6x. An exemplary protocol is provided in Example 8 for producing an exemplary construct suitable for homologous recombination. Figure 4 illustrates a preferred method for preparing an insertion construct to inactivate the *pbc*A gene in *S. pneumoniae*. In these methods, a gene encoding PbcA (a non-native C3 binding protein) is introduced into a cell containing PbcA. The non-native C3 binding protein is homologous to the PbcA in the cell to facilitate homologous recombination and the production of insertional mutations. The presence of the insertion mutant was confirmed by Southern blot for a number of mutated R6X and CP1200 strains following transformation of these strains with the mutating construct pBV004 (see Figure 4). Mutated strains were tested by Southern blot using a 1.5 kb fragment of *pbc*A or vector pVA891 probes. Wild type strain BD23 was not transformed in these studies. Results of the Southern blot experiments are provided in Figure 5.

The insertion mutants were tested for their ability to produce PbcA protein using either human C3 or antibody to PbcA (Figure 4). Results indicated that no protein approximating the size of PbcA was detectable in supernatants from the *S. pneumoniae* strains tested on Western blot. Further, neither C3 nor antibody to PbcA bound on the Western blots. Therefore, the mutants are useful to assess functional aspects of PbcA and to serve as a negative control for a variety of PbcA-related experiments.

As noted above, PbcA binds noncovalently to C3. *S. pneumoniae* is generally cleared from the body by the covalent opsonic deposition of C3b on the pneumococcal capsule or cell wall, followed by phagocytosis via C3 receptors or Fc receptors on neutrophils or monocytes/macrophages. Nonopsonic (i.e. non-covalent) binding of C3 by pneumococcal surface proteins,

such as PbcA, suggests a mechanism whereby pneumococci can evade opsonization. Without intending to limit the scope of this invention, it is possible that PbcA protein from *S. pneumoniae* could bind C3 *in vivo* and reduce the amount of C3 available for opsonization. Antibodies to PbcA could block the C3-binding effect and restore the opsonic activity of C3 in plasma.

PbcA can be used in *in vitro* assays to assess the effect of PbcA on cells. For example, in Example 6, purified PbcA was added to cell cultures to study the role of PbcA in *S. pneumoniae* pathogenesis. PbcA can be tested for its toxicity on a variety of cells as well as tested in *in vivo* models for toxicity. The results of Example 6 indicated that PbcA, as an intact protein, was toxic to pulmonary epithelial cells and that it stimulated production of the cytokine IL-8 from the epithelial cells. IL-8 is a cytokine that, among other things, stimulates neutrophil migration. Increased concentrations of neutrophils are observed in the lungs of patients with significant *S. pneumoniae* infection in the lung passageways. Neutrophils and other white blood cells produce a variety of degradatory enzymes that damage lung tissue in *S. pneumoniae* infection and, based on these studies, lung damage can be the result of enzymatic release from white blood cells, PbcA, or both.

S. pneumoniae can colonize the nasopharynx, infect the lung and ultimately disseminate to the blood, meninges or other sites. Antibody to PbcA can be tested for its ability to reduce the toxicity of the organism to the lung tissue and antibody to PbcA can be tested for its ability to prevent PbcA binding to C3 and to permit C3 to remain available during *S. pneumoniae* infection. Similarly, peptides and polypeptides to PbcA can be administered to mammals and used in studies to assess the ability of the immune system to produce antibody to limit *S. pneumoniae* infection.

Animal models to study *S. pneumoniae* pathogenesis are known in the art. These include, but are not limited to, the chinchilla model for otitis media, the infant rat model for *S. pneumoniae* colonization and bacteremia, the mouse model for colonization, bacteremia and meningeal infection and a rabbit model for studying infection in the meninges. Those of ordinary skill in the art

will recognize that antibodies to PbcA (whether exogenously administered or the product of immunization with all or a part of PbcA) can be tested in these models for their ability to limit or inhibit *S. pneumoniae* infection. Exogenously (i.e., passively) administered antibody can be given through a variety of parenteral routes including, but not limited to, intravenous administration or administration to the air passages of a mammal such as a mouse, chinchilla, rat, rabbit or human.

All references and publications cited herein are expressly incorporated by reference into this disclosure. Particular embodiments of this invention will be discussed in detail and reference has been made to possible variations within the scope of this invention. There are a variety of alternative techniques and procedures available to those of skill in the art which would similarly permit one to successfully perform the intended invention.

Example 1
Identification of PbcA from *S. pneumoniae*

10 ml of *S. pneumoniae* strain CP1200 (obtained from D.A. Morrison, University of Illinois, Champagne-Urbana, Illinois and described in Havarstein LF, et al. *Proc. Natl. Acad. Sci. (USA)* 1995;92:11140-11144) was grown to exponential phase ($O.D._{620} \cong 0.3$) in Todd Hewitt broth (Fisher, Pittsburgh, PA) or in a synthetic medium ($O.D._{620} \cong 0.15$, media described by Sicard, A.M. *Genetics* 1964 59:31-44). Pneumococcal cells were pelleted and the supernatant was removed and precipitated with a final concentration of 10% trichloroacetic acid (TCA) at 4°C overnight and samples were electrophoresed on 15% sodium dodecyl sulfate polyacrylamide gels (SDS-PAGE) under non-reducing conditions. Pneumococcal proteins separated by electrophoresis were then transferred to nitrocellulose for Western blotting. After blocking of the Western blot according to standard protocols (Harlow , et al. *Antibodies; A Laboratory manual*. Cold Spring Harbor, NY; Cold Spring harbor Laboratory Press, 1988; 471-510) the blot was incubated with 10 mls of binding buffer containing about 2 µg/ml of purified human C3 (Hostetter et al. *J. Infect. Dis.*

1994; 150:653-661), labeled with biotin. The blot was washed and incubated with a 1:20,000 dilution of HRP-avidin for 60 minutes at room temperature and developed using the Supersignal™ system (Pierce, Rockford, IL) according to manufacturer's instructions. Purified human C3 bound to a band of about 90 kDa under non-reducing and reducing conditions on a 15% SDS-PAGE gel. Similarly, a protein of about 90 kDa was detected when cells were lysed in 5% SDS at room temperature for 30 minutes and following centrifugation, the supernatant was separated on a 15% SDS-PAGE gel.

Experiments were repeated with the substitution of 2 μg/ml biotinylated C3 after treatment with methylamine (using the methods disclosed in Hostetter, et al. *J. Infect. Dis.* 1984; 150:653-661) to demonstrate that the 90 kDa protein in pneumococcal supernatants and lysates could bind non-opsonic forms of C3. Results again indicated that a 90 kDa protein was recognized by C3. A variety of pneumococcal strains were also tested. C3 bound to PbcA identified in Western blots using supernatants from a variety of virulent pneumococcal strains (serotypes 1,3 [4 strains], 4, 14, 19F)

Example 2
Purification of PbcA

S. pneumoniae CP1200 was grown to mid-exponential phase in 4 liters of Todd Hewitt broth at 37°C. Pneumococcal cells were pelleted by centrifugation at 10,000 x g for 10 minutes. Proteins in the supernatant were precipitated with a final concentration of 10% trichloroacetic acid at 4 °C overnight. The precipitate was resuspended in 40 mls of 100 mM Tris and the final pH adjusted to 7.0. The resuspended proteins were chromatographed on a 1.2 ml column of Thiopropyl Separaose 6B coupled by a disulfide bond to 4 mg methylamine treated human C3. The column was then washed with 40 ml of 100 mM Tris-HCl, pH 7.0 containing 0.5 M NaCl. PbcA was eluted from the column with 20% ethanol in the Tris-HCl/NaCl wash buffer. 1 ml. fractions were collected and analyzed by SDS-PAGE and C3 binding assay. A protein of about 90 kDa

to about 110 kDa (+/- 5 kDa) eluted from the C3 affinity column in fractions 2-10.

Multiple eluates from sequential purifications of PbcA from strain 1200 were pooled by precipitation with 90% ethanol to obtain sufficient sample for sequencing and further protein studies. Approximately 80 picomoles of the protein were subjected to amino terminal analysis and tryptic digestion for internal peptide sequencing at the Harvard Microchemical Facility (Cambridge, MA).

Example 3
Immunogenicity of PbcA

Immunogenicity of PbcA was assessed by growing 10 ml of *S. pneumoniae* strain CP1200 to exponential phase in Todd Hewitt broth, pelleting the cells, and precipitating proteins from the supernatant with 10% TCA overnight at 4°C. The next day, supernatant proteins were electrophoresed on 15% SDS-PAGE, transferred to nitrocellulose, blocked with skim milk in a standard protocol (Harlow, et al. *supra*) and incubated with a 1:10,000 dilution of acute or convalescent serum from a patient infected with *S. pneumoniae* (Dr. E. Janoff, Minneapolis VA Hospital, Minneapolis, MN). The blot was washed according to methods disclosed in Harlow et al. and incubated with a 1:50,000 dilution of commercially purchased goat anti-human IgG conjugated to horseradish peroxidase (Chemicon, Temicula, CA). The blot was washed and developed with the Supersignal™ system according to manufacturer's instructions.

Western blots from these studies demonstrated that acute serum did not contain IgG antibodies to any pneumococcal proteins released into the supernatant but that convalescent serum contained IgG antibodies that recognized a protein of 90 kDa, consistent with the mass of PbcA. These experiments confirmed that PbcA elicited an immune response in humans recovering from *S. pneumoniae* infection and indicated that PbcA is recognized by the human immune system.

Example 4
PbcA is implicated in *S. pneumoniae* virulence

Virulent pneumococcal isolates typically cause otitis media and influx of leukocytes after inoculation in concentrations as low as 1×10^2 (Giebink et al. *J. Infect. Dis.* 1993; 167:347-355). Giebink et al. reported that an avirulent serotype, type 23F, was inoculated into the ears of chinchillas and that inoculum at concentrations of less than about 1×10^7 cells was unable to cause otitis media or inflammation.

Both the type 23F avirulent strain (GD 23, *supra*) and a type 23F virulent strain (BD23, Dr. Steve Pelton, Boston City Hospital, Boston, MA) were grown to mid-exponential phase in Todd Hewitt broth, the cells were pelleted, and supernatant proteins were precipitated in a final TCA concentration of 10% overnight at 4°C. The following day, the precipitate was resuspended in about 1 ml Tris, neutralized to pH 7.0, electrophoresed on 15% SDS-PAGE and then transferred to nitrocellulose. Incubation of the nitrocellulose membrane with 2 µg/ml of biotinylated, methylamine-treated C3 in binding buffer and development of the Western blot with avidin conjugated to horseradish peroxidase detected a PbcA band from a 15% SDS-PAGE gel in both cell lysates and supernatants from the virulent type 23F, but PbcA was completely absent in cell lysates and supernatants from the avirulent type 23F. In place of PbcA, a smaller band of 33 kDa was identified. This band may represent a degradation product or a truncated version of PbcA.

Example 5
Isolation of nucleic acid encoding PbcA

Degenerate oligonucleotides were obtained from a commercial supplier based on the sequence of SEQ ID NO:1 and SEQ ID NO:2. The oligonucleotides were used to amplify a 1500 bp sequence from CP1200 genomic DNA as template using standard polymerase chain reaction technology. Template, primers, and buffer were added for one 5-minute cycle at 94°C. Then dNTP's and Taq polymerase were added for 30 cycles, as follows: Double

stranded DNA was denatured for 1 min at 94°C, annealed for 1 min at 50°C, and extended for 2 min at 72°C. Final extension was completed in one 8-minute cycle at 72°C. The 1500 bp sequence was random primer labeled using a commercial kit and the sequence was used to screen the CP1200 genomic DNA library prepared under our direction by Stratagene (LaJolla, CA). Hybridization was performed under at least moderate stringency conditions and a variety of hybridization methods are provided in Sambrook et al. (*Molecular Cloning, A Laboratory Manual*, 1989 Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY). Exemplary hybridization conditions used were 65°C hybridization in 6X SSC (1M NaCl) using 5X Denhardt's, 20 mM sodium phosphate, 0.5% SDS and 100 µg/ml denatured sonicated salmon sperm DNA.

Three clones were identified that hybridized to the 1500 bp genomic DNA fragment; two overlapping clones of 4.3 and 5.3 kb, respectively and a third clone of 2.5 kb. Restriction mapping and hybridization studies suggested that the 5.3 kb clone contained an open reading frame encompassing the oligonucleotides from SEQ ID NO:1 and SEQ ID NO:2. A 3.2 kb fragment remaining after HindIII digestion and religation of the 5.3 kb clone is sequenced.

Example 6
Stimulation of IL-8 Production from Pulmonary Epithelial Cells in Response to PbcA

Monolayers of type II pulmonary epithelial cells (A549, American Type Culture Collection (ATCC), Rockville, MD) were incubated for 4 hours in culture supernatant (at least about 15µl supernatant diluted in 1 ml. serum free media (50% Hams F12/50% PBS) from exponentially growing *S. pneumoniae* (Pn) laboratory strain (CP1200, *supra)* or clinical isolates grown in Todd Hewitt broth (*supra*). Incubation of the epithelial cell monolayer in the culture supernatant induced the release of about 1169 pg/ml IL-8 while incubation in media alone without culture supernatant and incubation in pneumococcal growth medium alone without culture supernatant was significantly less ($p < 0.0001$, see table below).

Supernatants from a virulent *S. pneumoniae* clinical isolate, 23F, were also effective at stimulating IL-8 production from pulmonary epithelial cells (about 1,495 pg/ml IL-8). Supernatants from a 23F strain that did not produce otitis media in an animal model did not induce IL-8 release *in vitro*.

Increasing the time for which CP 1200 supernatants were incubated with the epithelial monolayer from 4 to 24 hours resulted in an increase in IL-8 production to approximately two times the level of IL-8 obtained after a 4 hour incubation.

SDS-PAGE analysis of proteins in CP1200 *S. pneumoniae* supernatants indicated that there a variety of proteins in the supernatants including five discrete bands correlating to identifiable proteins from *S. pneumoniae* that had sizes of about 180 kDa, 90 kDa, 57 kDa 42 kDa and 24 kDa. All five proteins were present in supernatants from the virulent 23F strain during exponential growth, but a band of about 90 kDa to about 110 kDa was absent in supernatants from the avirulent 23F strain.

In identical assays, about 150-200 ng of affinity purified PbcA protein elicited at least about 1200 pg/ml IL-8 from pulmonary epithelial cells. This level was similar to that elicited using supernatants from CP1200 and virulent 23F strains after a 4-hour incubation. Thus, not only is PbcA potent in eliciting IL-8 from pulmonary epithelial cells.

The combined results from 11 separate assays for IL-8 production are provided below:

Stimulus	**IL-8 Release (pg/ml) mean ±S.E.**
Medium alone	580 ± 40
Pneumo supt. Avir 23F(GD23)	471 ±72
Pneumo supt. Vir23F(BD23)	1,138 ± 83
Pneumo supt. CP1200	1,169 ±121
200 ng PbcA(CP1200)	1,679 ± 113
150 ng PbcA(CP1200)*	1,201
*assayed once	

Rabbit polyclonal antibody prepared against PbcA was used to determine whether the toxicity and IL-8 production stimulated by PbcA was

inhibited by PbcA-specific antibody as confirmation that the effects observed in these studies were attributable to PbcA.

Example 7
PbcA Binds Human C3

C3/PbcA ELISA Binding Assay

Purified PbcA was bound to ELISA 8-well strips (Costar, Cambridge, MA) overnight in binding buffer at room temp at varying concentrations (.5 μg/well and 0 μg/well). Wells were blocked with 5% Milk, 0.05% Tween 20 in PBS and 0.02% azide for 2 hours, washed 3 times with 0.05% Tween 20 in PBS and C3 was added (5 μg/well) in antibody diluent buffer (1% BSA, 1% Tween 20 in PBS) for 2 hours at 37°C. Plates were washed 3 times and incubated with HRP-conjugated goat anti-human C3 (1μg/ml) for 1 hour at 37°C. Plates were washed 3 times, developed with OPD (Zymed protocol) for 30 minutes, and the absorbence was read at A_{490} on an ELISA plate reader.

A standard curve was prepared to assess saturation of PbcA on the plate. Serial dilutions of PbcA protein were added (from 500 ng/well to 0 ng/well) to microtiter wells. Plates were incubated for 1 hour at 37°C with anti-PbcA (1:100), washed, incubated with HRP-conjugated goat anti-rabbit (1:2000) for 1 hour at 37°C. Plates were washed and developed with a 1:10 dilution of OPD developing buffer for 20 minutes at room temperature as per manufacturer's instructions and read at A_{490} on an ELISA plate reader.

Binding of C3 was determined by the immobilization of 500 ng PbcA/well. C3 was added in serial dilutions from 10 μg/well to 0 μg/well and incubated with immobilized PbcA for 2 hours at 37°C. Goat HRP-conjugated anti-human C3 (1 μg/well) was added and the absorbance was assessed. Figure 2 shows a dose-response curve for the binding of increasing amounts of C3 to constant amounts of immobilized PbcA.

Results indicated (see Figure 2) that increasing concentrations of C3 bound to 0.5 μg/well PbcA on the plate and saturated with a plateau of 6μg/well.

Initial blocking experiments were done using the IgG fraction of anti-PbcA polyclonal antibodies (2 to 8-fold molar excess over 500 mg PbcA added per well). The IgG fraction of serum from unimmunized rabbits served as a control. After PbcA protein was bound to the plate overnight and blocked for 2 hours, anti-PbcA antibodies were added for 1 hour at 37°C prior to the addition of C3. Plates were washed, 6 μg C3 was added to each well for 2 hours at 37°C. Plates were then incubated with goat HRP-conjugated anti-human C3 for 1 hour at 37°C then washed and developed for 4 minutes.

Both the IgG fraction of anti-PbcA antibodies and affinity purified antibodies inhibited C3 binding up to 50% while control IgG antibodies did not inhibit C3 binding. Figure 3a (Western blot) demonstrates that the IgG fraction of anti-PbcA blocked the binding of human C3 to PbcA immobilized on nitrocellulose. Figure 3b demonstrates that affinity-purified anti-PbcA blocked the binding of C3 to PbcA by 40-50%, while those antibodies remaining in rabbit serum after the removal of anti-PbcA antibodies had no blocking effect.

Example 8
Production of *pbcA* Insertion/Duplication Mutants of *S. Pneumoniae*

Generation of a *pbcA* insertion/duplication construct

To generate an insertion/duplication mutant of *pbcA*, a 761 bp fragment was isolated from pDF145 (a plasmid containing *pbcA*) first by digestion with ApoI (New England Biolabs) followed by the addition of DNA Polymerase I, Large (Klenow) Fragment to create a blunt end, and second by digestion with NheI to create a 5' overhang sticky end compatible with XbaI in the vector pVA891. This vector is a streptococcal/*E. coli* shuttle vector that replicated in *E. coli* but not in *S. pneumoniae* (from Dr. Gary Dunny, Dept. Microbiology, University of Minnesota, Minneapolis, MN and described in Macrina FL, et al. *Gene* 25:145-150, 1983). The 761 bp fragment was isolated on

a 0.7% agarose gel. Concurrently, the vector pVA891 was digested with ClaI and treated with Klenow to create a blunt end, followed by digestion with XbaI to create a 5' sticky end. The 761 bp fragment of *pbc*A was ligated into the pneumococcal vector pVA891 and transformed into competent DH5α *E. coli*. Tranformants were selected by resistance to chloramphenicol and plasmid DNA from 12 clones was cut with EcoRI and SpeI (the 761 fragment introduced a SpeI site). This confirmed the presence of the insert DNA in clones 2-12 and the absence of the insert in the control vector clone. Clone 4 was selected as construct pBV004 and DNA was transformed into unencapsulated laboratory strains R6x and CP1200 and selected by erythromycin (Erm) resistance.

Transformation of Knockout Construct for Insertion/Duplication Mutagenesis

Pneumococcal strains CP1200 and R6x were grown to OD_{550}= 0.2 and stored as frozen stock aliquots. Pneumococcal cultures were diluted 1:100 to 0.002 and grown to OD_{550}= 0.02. Competence was induced by the Morrison CSP protocol (Haverstein, L. et al. *Proc. Natl. Acad. Sci.* (USA) 92:11140-11144, 1995). To 100µl of cells, competence stimulating peptide (CSP) 100ng, was added along with 500ng to about 1 µg of construct pBV004 DNA. Cells were incubated for 30-40 minutes in a 37°C water bath with aeration to maintain constant temperature. After transformation, cells were diluted 1:10 in one ml total volume of THB+Y. DNase I (10µg/ml) was added and cells were incubated an additional 90 minutes to allow integration. The transformation mixture was diluted 1:10 and 100µl of cells were plated in a 4-layer agar overlay procedure as follows: first overlay, 3 mls THB agar; second overlay, 1.5 mls THBY + 1.5 mls THB agar + 100 µl transformation mixture, incubate 1 hour at 37° C; third overlay, 3 mls THB agar; fourth overlay, 3 mls THB agar + .05µg/ml Erythromycin.

Genomic DNA was isolated from wild type R6x, CP1200 and BD23(virulent 23F strain *supra*) strains as well as from insertional mutants in R6x and CP1200. Genomic DNA was digested with EcoRI (which does not

cut within the *pbc*A gene) and electrophoresed in 0.7% agarose, blotted onto nylon membrane and Southern blot performed with digoxygenin-labeled probes.

Referring to Figure 5, on the left are the results of a the Southern blot using the 1.5 kb fragment of *pbc*A as the probe. The probe hybridized with wild-type pneumococcal DNA in a single band and hybridized with two bands in the mutants R62r, R63r, R6x5 and CP1r, as expected.

On the right, the same blot was stripped and reprobed with pVA891 vector DNA. Wild-type DNA does not hybridize with pVA891, while the vector probe hybridizes at the same two bands indicating that the insertion mutagenesis was within the *pbc*A gene.

To confirm that the insertion mutated the gene and disrupted protein production, a Western blot analysis was performed (see Figure 6).

On the left, trichloroacetic acid (TCA) precipitated supernatants from strain R6x and the R62r mutant were electrophoresed on a 7.5% SDS-PAGE gel and blotted with either anti-PbcA antibodies and HRP-goat- anti-rabbit antibodies for detection of PbcA or with biotinylated C3 and HRP-avidin for detection of C3 binding. Insertional mutants did not synthesize PbcA and failed to bind to anti-PbcA antibodies or biotinylated C3. Absence of protein and lack of C3 binding was seen in the mutant compared to wild-type. The PbcA protein from R6x is larger than that of CP1200.

The original CP1200 detected a 90 kDa protein and had a truncated choline binding region due to premature termination. The R6x PbcA protein was about 105 kDa and appears to contain at least 9 choline binding repeats.

It will be appreciated by those skilled in the art that while the invention has been described above in connection with particular embodiments and examples, the invention is not necessarily so limited and that numerous other embodiments, examples, uses, modifications and departures from the embodiments, examples and uses may be made without departing from the inventive scope of this application.

TABLE 1

DNA SEQ OF *pbcA* GENE (SEQ ID NO:4)

1 GTAATACGAC TCACTATAGG GCGAATTGGG TACCGGGCCC CCCCTCGAGG

51 TCGACGGTAT CGATAAGCTT ATGCTTGTCA ATAATCACAA ATATGTAGAT

101 CATATCTTGT TTAGGACAGT AAAACATCCT AATTACTTTT TAAATATTCT

151 TCCTGAGTTG ATTGGCTTGA CCTTGTTGAG TCATGCTTAT GTGACTTTTG

201 TTTTAGTTTT TCCAGTTTAT GCAGTTATTT TGTATCGACG AATAGCTGAA

251 GAGGAAAAGC TATTACATGA AGTTATAATC CCAAATGGAA GCATAAAGAG

301 ATAAATACAA AATTCGATTT ATATACAGTT CATATTGAAG TAATATAGTA

351 AGGTTAAAGA AAAAATATAG AAGGAAATAA ACATGTTTGC ATCAAAAAGC

401 GAAAGAAAAG TACATTATTC AATTCGTAAA TTTAGTATTG GAGTAGCTAG

451 TGTAGCTGTT GCCAGTCTTG TTATGGGAAG TGTGGTTCAT GCGACAGAGA

501 ACGAGGGAAG TACCCAAGCA GCCACTTCTT CTAATATGGC AAAGACAGAA

551 CATAGGAAAG CTGCTAAACA AGTCGTCGAT GAATATATAG AAAAAATGTT

601 GAGGGAGATT CAACTAGATA GAAGAAAACA TACCCAAAAT GTCGCCTTAA

651 ACATAAAGTT GAGCGCAATT AAAACGAAGT ATTTGCGTGA ATTAAATGTT

701 TTAGAAGAGA AGTCGAAAGA TGAGTTGCCG TCAGAAATAA AAGCAAAGTT

751 AGACGCAGCT TTTGAGAAGT TTAAAAAAGA TACATTGAAA CCAGGAGAAA

801 AGGTAGCAGA AGCTAAGAAG AAGGTTGAAG AAGCTAAGAA AAAAGCCGAG

851 GATCAAAAAG AAGAAGATCG TCGTAACTAC CCAACCAATA CTTACAAAAC

901 GCTTGAACTT GAAATTGCTG AGTTCGATGT GAAAGTTAAA GAAGCGGAGC

951 TTGAACTAGT AAAAGAGGAA GCTAAAGAAT CTCGAAACGA GGGCACAATT

1001 AAGCAAGCAA AAGAGAAAGT TGAGAGTAAA AAAGCTGAGG CTACAAGGTT

1051 AGAAAACATC AAGACAGATC GTAAAAAAGC AGAAGAAGAA GCTAAACGAA

1101 AAGCAGATGC TAAGTTGAAG GAAGCTAATG TAGCGACTTC AGATCAAGGT

TABLE 1 (CONTD.)

DNA SEQ OF *pbcA* GENE (SEQ ID NO:4)

1151 AAACCAAAGG GGCGGGCAAA ACGAGGAGTT CCTGGAGAGC TAGCAACACC

1201 TGATAAAAAA GAAAATGATG CGAAGTCTTC AGATTCTAGC GTAGGTGAAG

1251 AAACTCTTCC AAGCTCATCC CTGAAATCAG GAAAAAAGGT AGCAGAAGCT

1301 GAGAAGAAGG TTGAAGAAGC TGAGAAAAAA GCCAAGGATC AAAAAGAAGA

1351 AGATCGCCGT AACTACCCAA CCAATACTTA CAAAACGCTT GACCTTGAAA

1401 TTGCTGAGTC CGATGTGAAA GTTAAAGAAG CGGAGCTTGA ACTAGTAAAA

1451 GAGGAAGCTA AGGAACCTCG AGACGAGGAA AAAATTAAGC AAGCAAAAGC

1501 GAAAGTTGAG AGTAAAAAAG CTGAGGCTAC AAGGTTAGAA AACATCAAGA

1551 CAGATCGTAA AAAAGCAGAA GAAGAAGCTA AACGAAAAGC AGCAGAAGAA

1601 GATAAAGTTA AAGAAAAACC AGCTGAACAA CCACAACCAG CGCCGGCTAC

1651 TCAACCAGAA AAACCAGCTC CAAAACCAGA GAAGCCAGCT GAACAACCAA

1701 AAGCAGAAAA AACAGATGAT CAACAAGCTG AAGAAGACTA TGCTCGTAGA

1751 TCAGAAGAAG AATATAATCG CTTGACTCAA CAGCAACCGC CAAAAACTGA

1801 AAAACCAGCA CAACCATCTA CTCCAAAAAC AGGCTGGAAA CAAGAAAACG

1851 GTATGTGGTA CTTCTACAAT ACTGATGGTT CAATGGCAAC AGGATGGCTC

1901 CAAAACAACG GTTCATGGTA CTATCTAAAC GCTAATGGTG CTATGGCGAC

1951 AGGATGGCTC CAAAACAATG GTTCATGGTA CTATCTAAAC GCTAATGGTT

2001 CAATGGCAAC AGGATGGCTC CAAAACAATG GTTCATGGTA CTACCTAAAC

2051 GCTAATGGTG CTATGGCGAC AGGATAGCTC CAATACAATG GTTCATGGTA

2101 CTACCTAAAC AGCAATGGCG CTATGGCGAC AGGATGGCTC CAATACAATG

2151 GCTCATGGTA CTACCTCAAC GCTAATGGTG ATATGGCGAC AGGATGGCTC

2201 CAAAACAACG GTTCATGGTA CTACCTCAAC GCTAATGGTG ATATGGCGAC

2251 AGGATGGCTC CAATACAACG GTTCATGGTA TTACCTCAAC GCTAATGGTG

2301 ATATGGCGAC AGGTTGGGTG AAAGATGGAG ATACCTGGTA CTATCTTGAA

2351 GCATCAGGTG CTATGAAAGC AAGCCAATGG TTCAAAGTAT CAGATAAATG

2401 GTACTATGTC AATGGCTCAG GTGCCCTTGC AGTCAACACA ACTGTAGATG

2451 GCTATGGAGT CAATGCCAAT GGTGAATGGG TAAACTAAAC CTAATATAAC

TABLE 1 (CONTD.)

DNA SEQ OF *pbcA* GENE (SEQ ID NO:4)

2501 TAGTTAATAC TGACTTCCTG TAAGAACTTT TTAAAGTATT CCCTACAAAT

2551 ACCATATCCT TTCAGTAGAT AATATACCCT TGTAGGAAGT TTAGATTAAA

2601 AAATAACTCT GTAATCTCTA GCCGGATTTA TAGCGCTAGA GACTACGGAG

2651 TTTTTTTGAT GAGGAAAGAA TGGCGGCATT CAAGAGACTC TTTAAGAGAG

2701 TTACGGGTTT TAAACTATTA AGCCTTCTCC AATTGCAAGA GGGCTTCAAT

2751 CTCTGCTAGG GTGCTAGCTT GCGAAATGGC TCCACGGAGT TTGGCAGCGC

2801 CAGATGTTCC ACGGAGATAG TGAGGAGCGA GGCCGCGGAA TTCACGAACT

2851 GCGACGTTTT CTCCTTTGAG GTTAATCAAT CGTTTCAGGA ATTCCGGAAT

2901 TCCGGAATTC CGGAATTCCG GAATTCCGGA ATTCCTGCAG CCCGGGGGAT

2951 CCACTAGTTC TAGAGCGGCC GCCACCGCGG TGGAGCTCCA GCTTTTGTTC

3001 CCTTTAGTGA GGGTTAATTT CGA

TABLE 2

*pbc*A DNA (SEQ ID NO:5)

PbcA PROTEIN (SEQ ID NO:6)

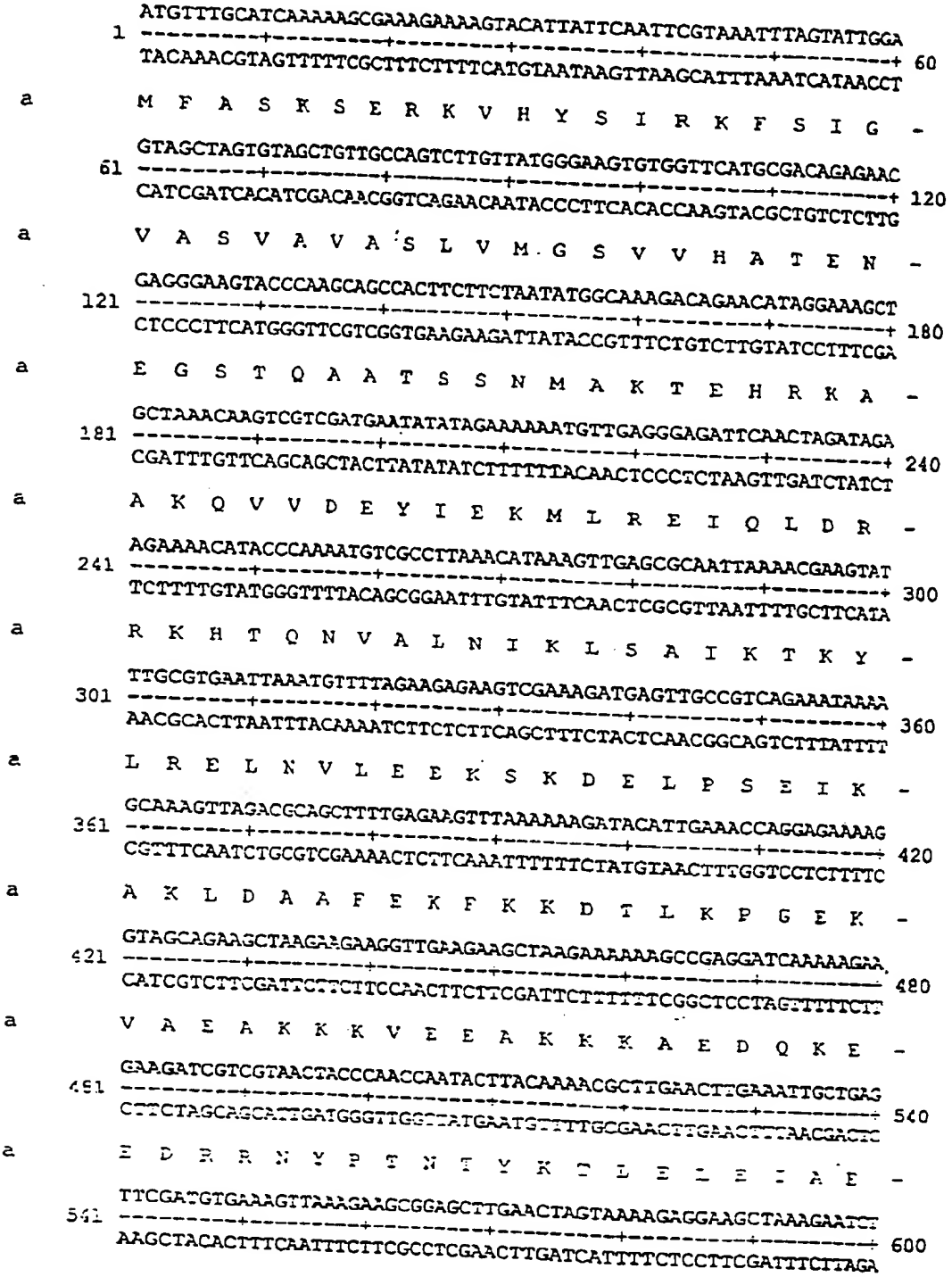


TABLE 2 (CONTD.)

*pbc*A DNA (SEQ ID NO:5)

PbcA PROTEIN (SEQ ID NO:6)

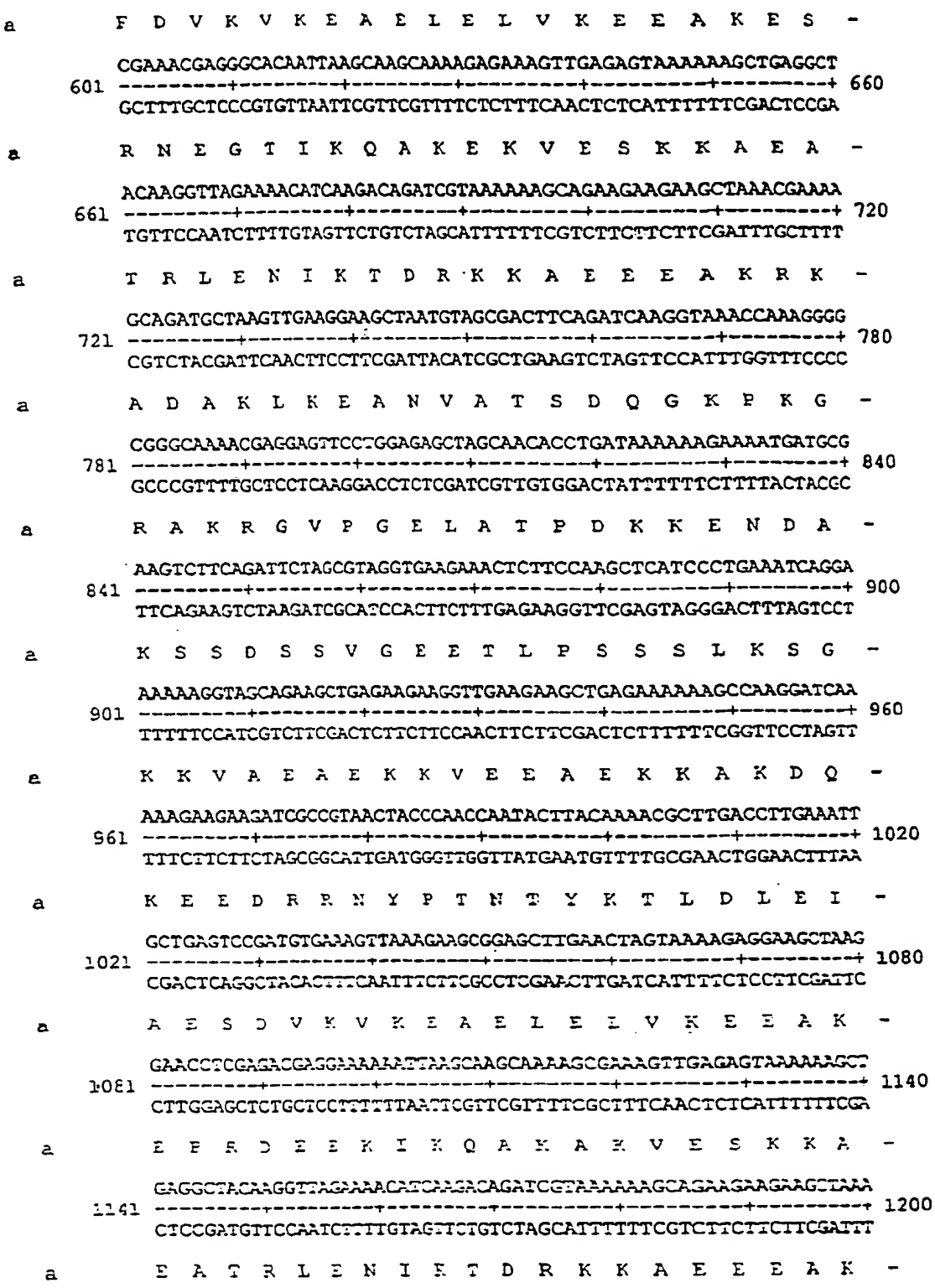


TABLE 2 (CONTD.)

*pbc*A DNA (SEQ ID NO:5)

PbcA PROTEIN (SEQ ID NO:6)

```
      CGAAAAGCAGCAGAAGAAGATAAAGTTAAAGAAAAACCAGCTGAACAACCACAACCAGCG
1201  ---------+---------+---------+---------+---------+---------+
      GCTTTTCGTCGTCTTCTTCTATTTCAATTTCTTTTTGGTCGACTTGTTGGTGTTGGTCGC  1260

a       R  K  A  A  E  E  D  K  V  K  E  K  P  A  E  Q  P  Q  P  A  -

      CCGGCTACTCAACCAGAAAAACCAGCTCCAAAACCAGAGAAGCCAGCTGAACAACCAAAA
1261  ---------+---------+---------+---------+---------+---------+
      GGCCGATGAGTTGGTCTTTTTGGTCGAGGTTTTGGTCTCTTCGGTCGACTTGTTGGTTTT  1320

a       P  A  T  Q  P  E  K  P  A  P  K  P  E  K  P  A  E  Q  P  K  -

      GCAGAAAAAACAGATGATCAACAAGCTGAAGAAGACTATGCTCGTAGATCAGAAGAAGAA
1321  ---------+---------+---------+---------+---------+---------+
      CGTCTTTTTTGTCTACTAGTTGTTCGACTTCTTCTGATACGAGCATCTAGTCTTCTTCTT  1380

a       A  E  K  T  D  D  Q  Q  A  E  E  D  Y  A  R  R  S  E  E  E  -

      TATAATCGCTTGACTCAACAGCAACCGCCAAAAACTGAAAAAACCAGCACAACCATCTACT
1381  ---------+---------+---------+---------+---------+---------+
      ATATTAGCGAACTGAGTTGTCGTTGGCGGTTTTTGACTTTTTGGTCGTGTTGGTAGATGA  1440

a       Y  N  R  L  T  Q  Q  Q  P  P  K  T  E  K  P  A  Q  P  S  T  -

      CCAAAAACAGGCTGGAAACAAGAAAACGGTATGTGGTACTTCTACAATACTGATGGTTCA
1441  ---------+---------+---------+---------+---------+---------+
      GGTTTTTGTCCGACCTTTGTTCTTTTGCCATACACCATGAAGATGTTATGACTACCAAGT  1500

a       P  K  T  G  W  K  Q  E  N  G  M  W  Y  F  Y  N  T  D  G  S  -

      ATGGCAACAGGATGGCTCCAAAACAACGGTTCATGGTACTATCTAAACGCTAATGGTGCT
1501  ---------+---------+---------+---------+---------+---------+
      TACCGTTGTCCTACCGAGGTTTTGTTGCCAAGTACCATGATAGATTTGCGATTACCACGA  1560

a       M  A  T  G  W  L  Q  N  N  G  S  W  Y  Y  L  N  A  N  G  A  -

      ATGGCGACAGGATGGCTCCAAAACAATGGTTCATGGTACTATCTAAACGCTAATGGTTCA
1561  ---------+---------+---------+---------+---------+---------+
      TACCGCTGTCCTACCGAGGTTTTGTTACCAAGTACCATGATAGATTTGCGATTACCAAGT  1620

a       M  A  T  G  W  L  Q  N  N  G  S  W  Y  Y  L  N  A  N  G  S  -

      ATGGCAACAGGATGGCTCCAAAACAATGGTTCATGGTACTACCTAAACGCTAATGGTGCT
1621  ---------+---------+---------+---------+---------+---------+
      TACCGTTGTCCTACCGAGGTTTTGTTACCAAGTACCATGATGGATTTGCGATTACCACGA  1680

a       M  A  T  G  W  L  Q  N  N  G  S  W  Y  Y  L  N  A  N  G  A  -

      ATGGCGACAGGATAG
1681  ---------+-----  1695
      TACCGCTGTCCTATC

a       M  A  T  G  *  -
```

WHAT IS CLAIMED IS:

1. An isolated protein comprising SEQ ID NO:1 and at least two choline binding repeats.

2. The protein of claim 1 wherein the protein is isolated from *S. pneumoniae*.

3. The protein of claim 1 wherein the protein binds human complement protein C3.

4. The protein of claim 1, wherein the protein is a recombinant protein.

5. The protein of claim 1 having a molecular weight as determined on a 15% polyacrylamide gel of between about 90 kDa to about 110 kDa.

6. The protein of claim 1 further comprising a proline rich region.

7. The protein of claim 1 further comprising SEQ ID NO:2 or SEQ ID NO:3.

8. An isolated protein comprising SEQ ID NO:6.

9. An isolated protein capable of binding to, but not cleaving or degrading, human complement C3 and wherein the protein comprises SEQ ID NO:1.

10. The protein of claim 9 wherein the protein is isolated from *S. pneumoniae*.

11. The protein of claim 9 further comprising a proline rich region.

12. The protein of claim 9 further comprising SEQ ID NO:2.

13. The protein of claim 9 having at least about 95% homology to a C3 binding protein from *S. pneumoniae*.

14. The protein of claim 9 having a molecular weight as determined on a 15% polyacrylamide gel of between about 90 kDa to about 110 kDa.

15. The protein of claim 9, wherein the protein is a recombinant protein.

16. A recombinant protein comprising SEQ ID NO:1, wherein the protein has a molecular weight as determined on a 15% polyacrylamide gel of between about 90 kDa to about 110 kDa.

17. The protein of claim 16 wherein the protein binds human complement protein C3.

18. The protein of claim 16 further comprising a proline rich region.

19. The protein of claim 17, wherein the protein does not cleave or degrade human complement protein C3.

20. A protein comprising amino acids 1-410 of SEQ ID NO:6.

21. A protein that binds, but does not cleave or degrade, human complement protein C3, wherein nucleic acid encoding the protein hybridizes to SEQ ID NO:4 under hybridization conditions of 6XSSC, 5X Denhardt, 0.5% SDS, and 100 µg/ml fragmented and denatured salmon sperm DNA hybridized overnight at 65°C and washed in 2X SSC, 0.1% SDS one time at room temperature for about 10 minutes followed by one time at, 65°C for about 15 minutes followed

by at least one wash in 0.2XSSC, 0.1% SDS at room temperature for at least 3-5 minutes.

22. The protein of claim 21 further comprising SEQ ID NO:1.

23. The protein of claim 21 further comprising SEQ ID NO:2 or SEQ ID NO:3.

24. The protein of claim 21 wherein the protein is at least 15 amino acids in length.

25. The protein of claim 21, wherein the protein is a recombinant protein.

26. The protein of claim 21 further comprising a proline rich region.

27. The protein of claim 21, wherein the protein is a synthetic peptide.

28. The protein of claim 21, wherein the protein is a peptide of at least 15 amino acids from SEQ ID NO:6.

29. Antibody capable of specifically binding to the protein of claim 21.

30. The antibody of claim 29, wherein the antibody is a monoclonal antibody.

31. The antibody of claim 29, wherein the antibody is a polyclonal antibody.

32. The antibody of 30, wherein the antibody is rodent-derived.

33. Nucleic acid encoding a protein comprising at least two choline binding domains and SEQ ID NO: 1.

34. The nucleic acid of claim 33 further comprising a proline rich region.

35. The nucleic acid of claim 33 isolated from an *S. pneumoniae* genome.

36. The nucleic acid of claim 33 capable of hybridizing to SEQ ID NO:4.

37. The nucleic acid of claim 33 wherein the protein encoded by the nucleic acid binds to human complement protein C3.

38. The nucleic acid of claim 33 in a nucleic acid vector.

39. The nucleic acid of claim 33 wherein the vector is an expression vector.

40. The nucleic acid of claim 38 in a cell.

41. The nucleic acid of claim 40 wherein the cell is a bacterium.

42. A bacterium expressing a recombinant protein according to claim 21.

43. The nucleic acid of claim 39 in a cell.

44. Protein produced by the nucleic acid of claim 43.

45. Isolated nucleic acid encoding a protein comprising SEQ ID NO:1 and a proline rich region wherein the protein encoded by the nucleic acid binds but does not cleave or degrade human complement C3.

46. An isolated nucleic acid fragment encoding an about 90 kDa to about 110 kDa protein with C3 binding activity, wherein the nucleic acid fragment has at least 80% homology to at least 500 bp from nucleic acids 1-1500 of SEQ ID NO:5.

47. An isolated nucleic acid fragment comprising base pairs 1-1500 of SEQ ID NO:5.

48. A method for isolating a C3 binding protein from a bacterium comprising the steps of:

obtaining a protein sample from a bacterium;

applying the sample to a solid support comprising methylamine treated complement protein C3;

washing the solid support; and

removing a C3 binding protein from the solid support in a solution comprising alcohol;

wherein the C3 binding protein does not cleave or degrade C3.

49. The method of claim 48 wherein the bacterium is *S. pneumoniae*.

50. The method of claim 48 wherein the bacterium is *E. coli*.

51. The method of claim 48 wherein the solid support comprises an affinity column.

52. The method of claim 48 wherein the alcohol is ethanol.

53. The method of claim 52 wherein the solution comprising alcohol is a buffer comprising 20% ethanol.

54. C3 binding protein preparable by the method of claim 48.

55. A method for producing an immune response to *S. pneumoniae* comprising the steps of:

administering a therapeutically effective amount of at least a portion of a protein to a mammal, wherein the protein binds but

does not cleave or degrade human complement protein C3 and, wherein nucleic acid encoding the protein hybridizes to SEQ ID NO:4 under hybridization conditions of 6XSSC, 5X Denhardt, 0.5% SDS, and 100 µg/ml fragmented and denatured salmon sperm DNA, hybridized overnight at 65°C and washed in 2x SSC, 0.1% SDS one time at room temperature for about 10 minutes followed by one time at 65°C for about 15 minutes followed by at least one wash in 0.2xSSC, 0.1% SDS at room temperature for at least 3-5 minutes; and
detecting an immune response to the protein.

56. The method of claim 55 wherein the protein is at least 15 amino acids in length.

57. The method of claim 55 wherein the protein is a chimeric protein.

58. The method of claim 55 wherein the protein comprises SEQ ID NO:1.

59. The method of claim 55 wherein the protein has a molecular weight on a 15% polyacrylamide gel of between about 90 kDa to about 110 kDa.

60. A method for reducing *S. pneumoniae* binding to C3 comprising the steps of:
administering a therapeutically effective amount of at least a portion of an antibody to a mammal, wherein the antibody specifically recognizes a C3 binding protein from *S. pneumoniae*, and wherein the C3 binding protein is a protein that binds, but does not cleave or degrade human complement protein C3.

61. The method of claim 60 wherein the antibody comprises at least one variable domain from a monoclonal antibody.

62. The method of claim 60 wherein the antibody is administered to the air passages of the mammal.

63. The method of claim 60 wherein the antibody is administered intravenously to a mammal.

64. A non-naturally occurring *S. pneumoniae* bacterium that does not express a detectable human complement C3 binding protein, wherein the C3 binding protein comprises SEQ ID NO:1.

65. The bacterium of claim 64 wherein the bacterium is the product of an insertion into the gene encoding the C3 binding protein.

66. The bacterium of claim 64 wherein the bacterium is a product of a deletion in the gene encoding the C3 binding protein.

67. The bacterium of claim 64 produced by homologous recombination of the gene encoding the C3 binding protein gene with at least a portion of a non-native C3 binding protein gene.

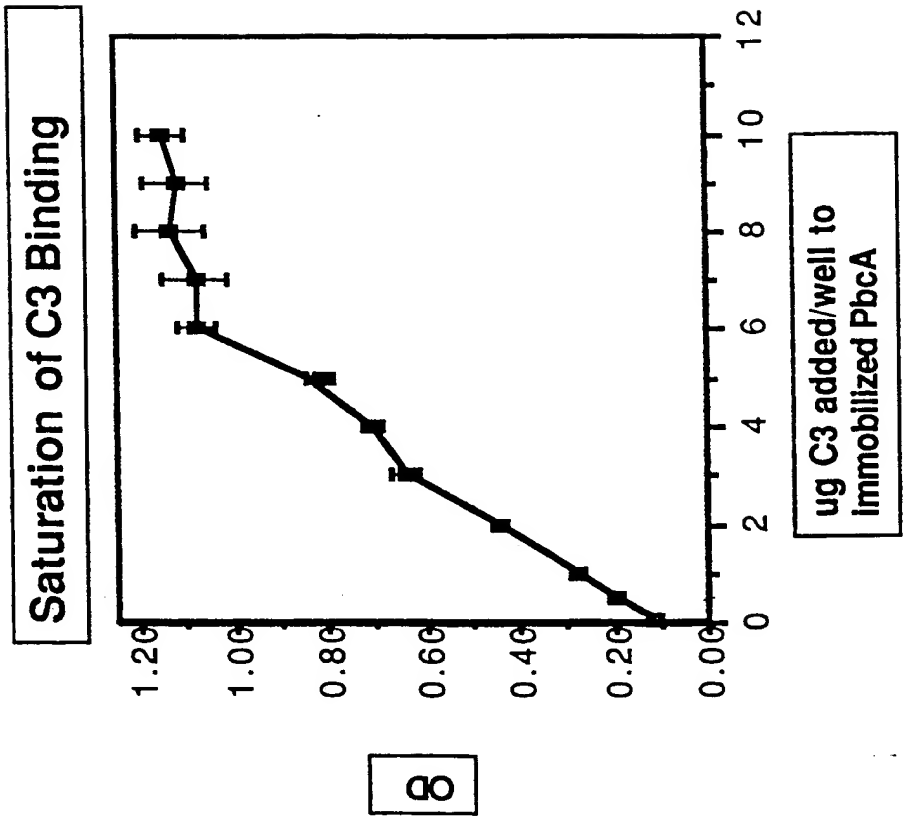
68. The method of claim 67 wherein the gene encoding at least a portion of the non-native C3 binding protein comprises a mutation within the non-native gene.

```
//

          1
1200      DRWKQENGMW YFYNTDGSMA TGWLQNNGSW YYLNANGAMA TGWLQNNGSW
r6x       ---KQENGMW YFYNTDGSMA TGWLQNNGSW YYLNANGAMA TGWLQNNGSW
23f       --LEYTRNGMW YFYNTDGSMA TGWLQNNGSW YYLNSNGAMA TGWLQNNGSW

          51
1200      YYLNANGSMA TGWLQNNGSW YYLNANGAMA TGWLQYNGSW YYLNSNGAMA
r6x       YYLNANGSMA TGWLQNNGSW YYLNANGAMA TGWLQYNGSW YYLNSNGAMA
23f       YYLNANGSMA TGWLQNNGSW YYLNANGSMA TGWLQYNGSW YYLNANGDMA

          101
1200      TGWLQYNGSW YYLNANGDMA TGWLQNNGSW YYLNANGDMA TGWLQYNGSW
r6x       TGWLQYNGSW YYLNANGDMA TGWLQNNGSW YYLNANGDMA TGWLQYNGSW
23f       TGWLQSNGSW YYLNANGDMA TGWLQNNGSW YYLNANGDMA TGWLQYNGSW

          151
1200      YYLNANGDMA TGWVKDGDTW YYLEASGAMK ASQWFKVSDK WYYVNGSGAL
r6x       YYLNANGDMA TGWVKDGDTW YYLEASGAMK ASQWFKVSDK WYYVNGSGAL
23f       YYLNANGDMF TGWVKDGDTW YYLEASGAMK ASQWFKVSDK WYYVNGSGAL

          201
                              223
1200      AVNTTVDGYG VNANGEW*TK H*Y
r6x       AVNTTVDGYG VNANGEW*TK P--
23f       AVNTTVDGYG VNANGEW*TK PNI
```

Fig. 1

Fig. 2

PbcA Binds C3 in ELISA



Anti-PbcA Inhibits Binding of C3 in Western and ELISA

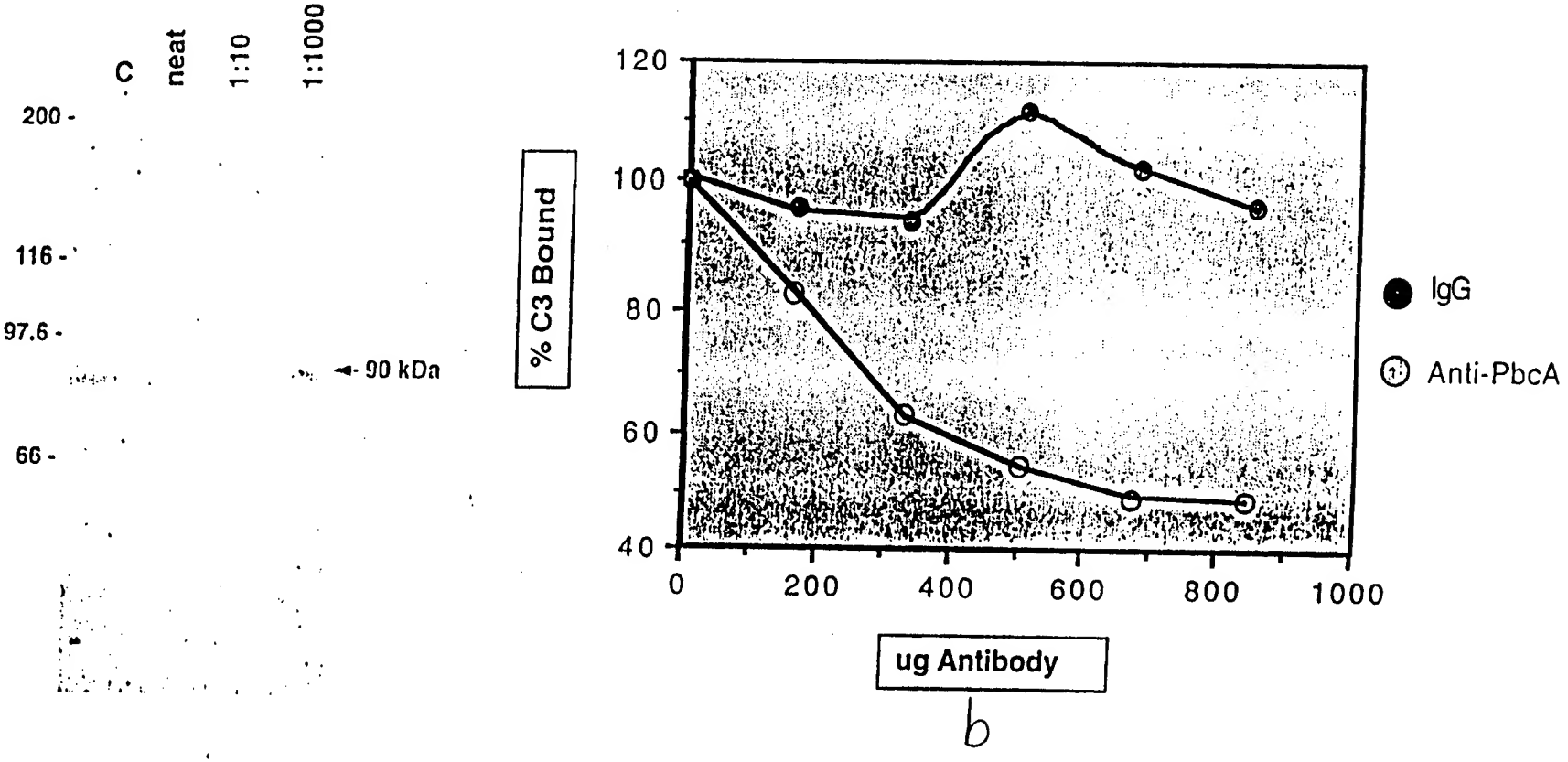
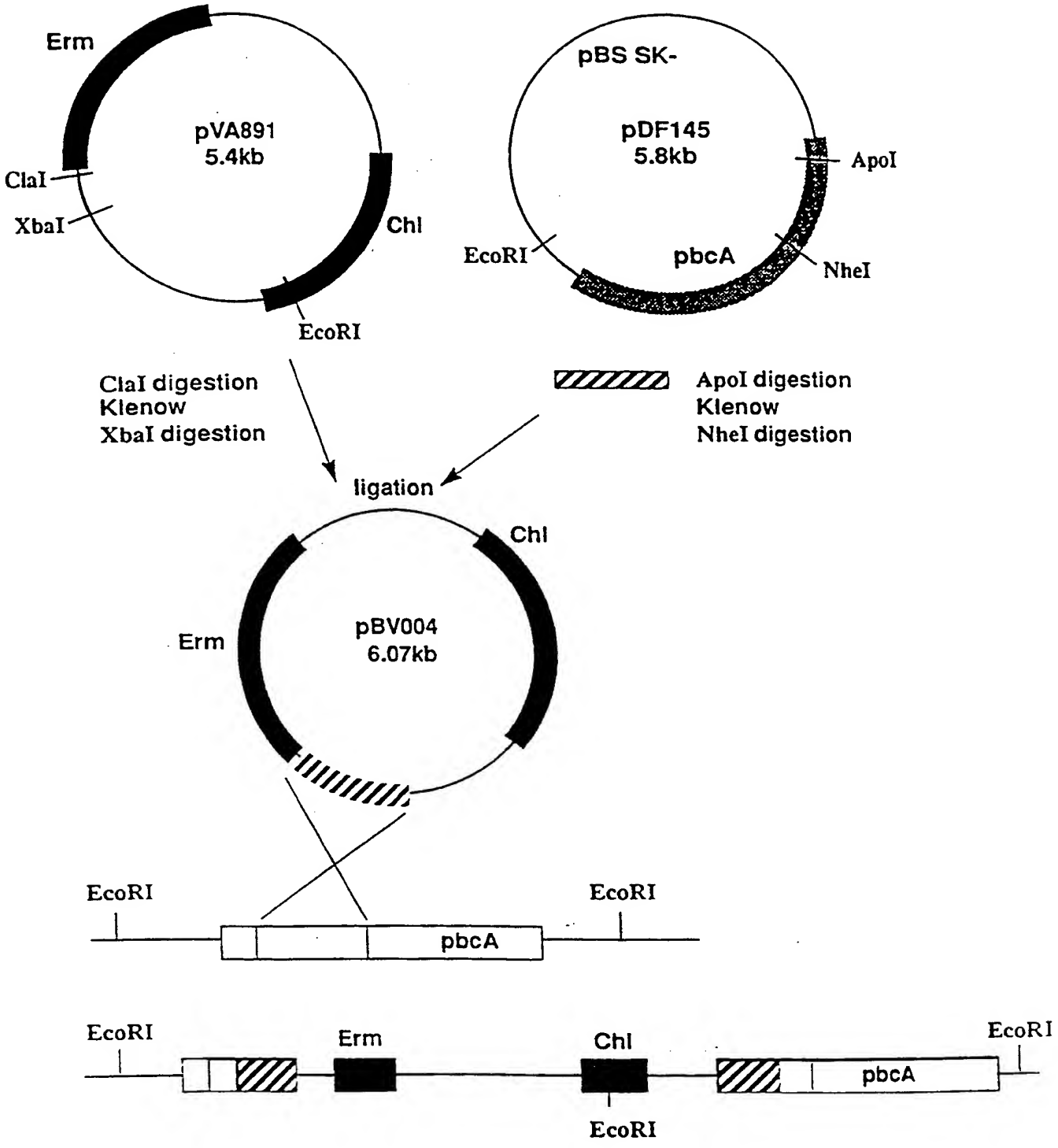


Fig. 3



Insertion-duplication mutagenesis

Fig. 4

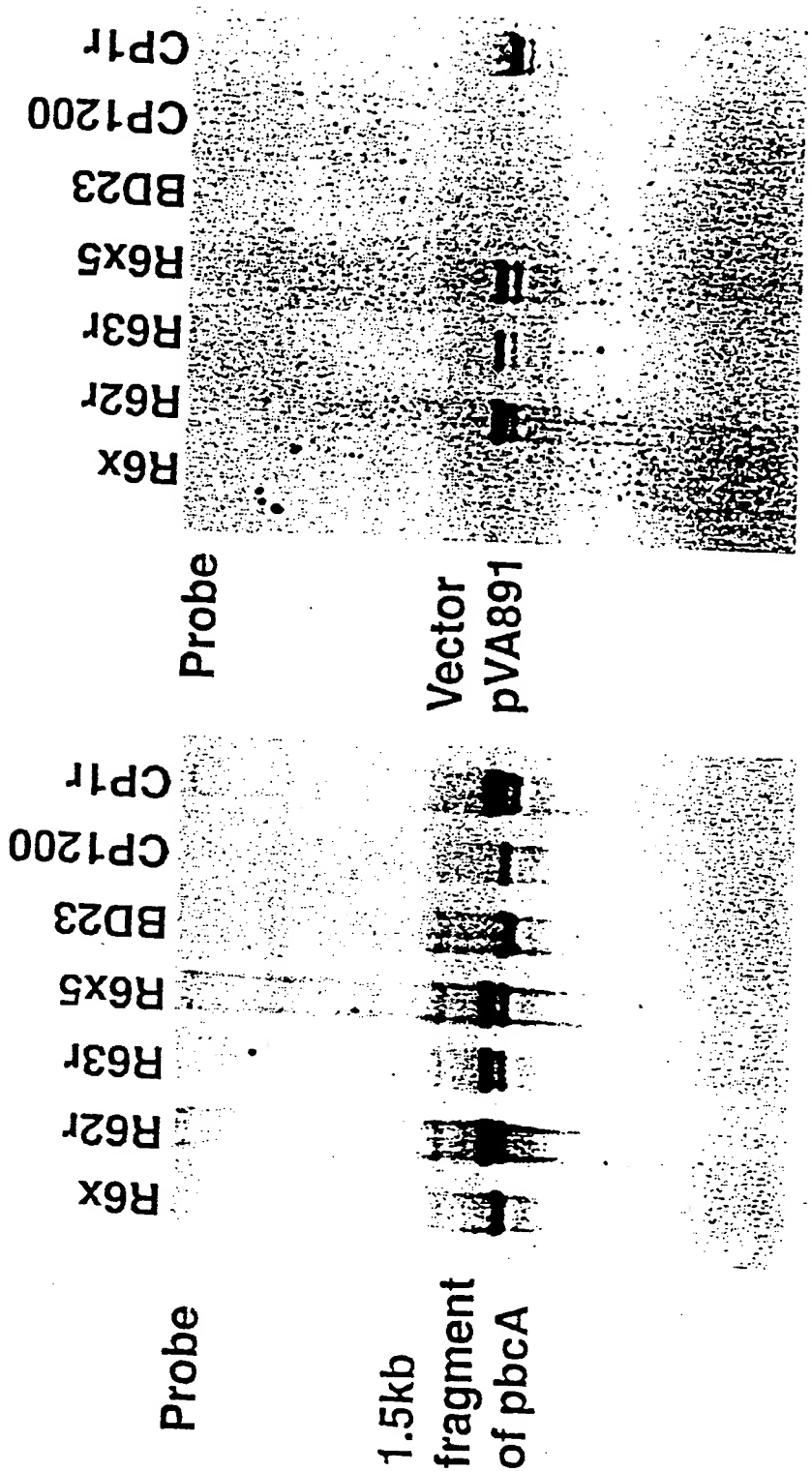


Fig. 5

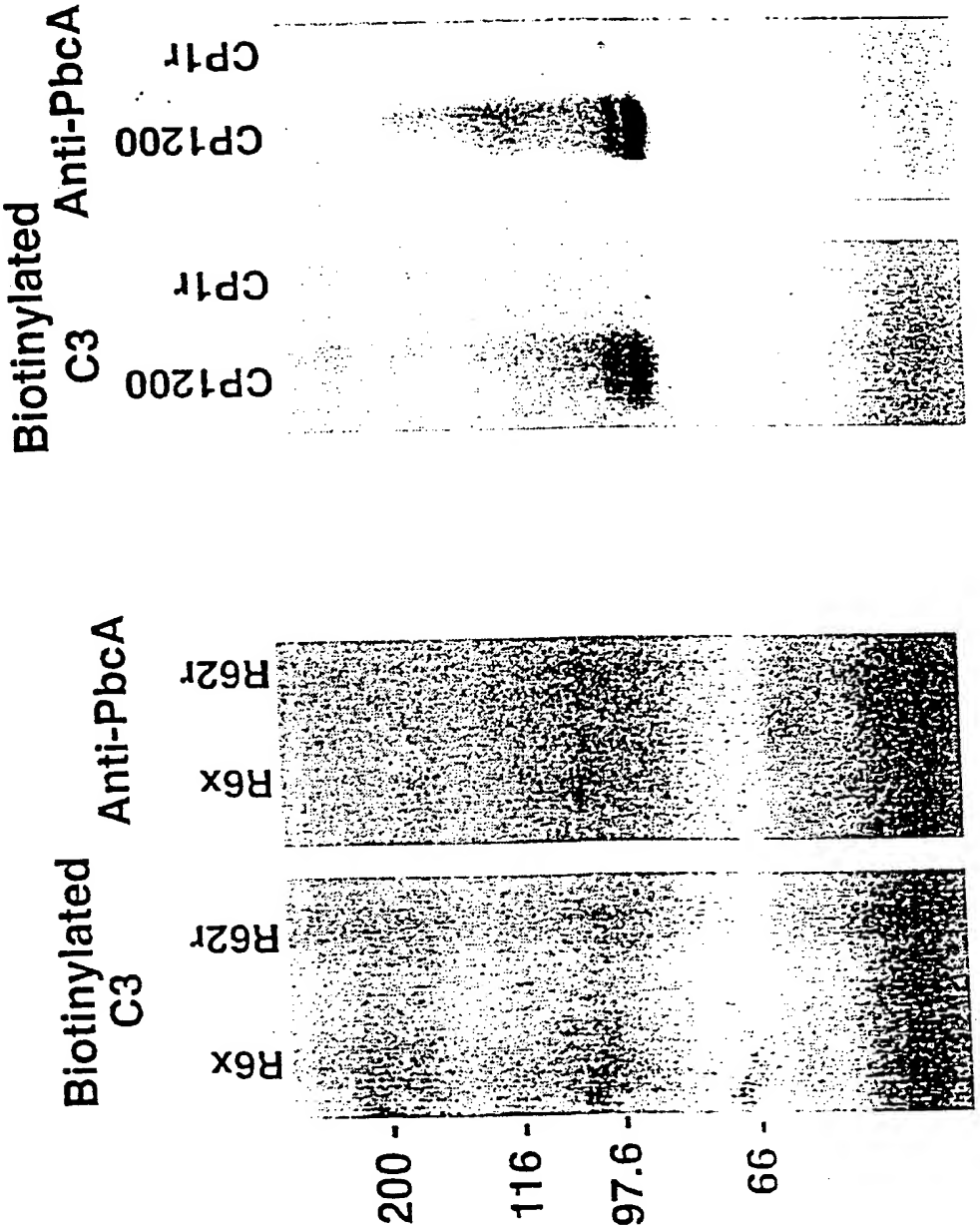


Fig. 6

PCT WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/31, 1/21, C07K 14/315, 16/12, A61K 39/09, 39/40 // (C12N 1/21, C12R 1:46)

A3

(11) International Publication Number: WO 98/21337

(43) International Publication Date: 22 May 1998 (22.05.98)

(21) International Application Number: PCT/US97/20586

(22) International Filing Date: 12 November 1997 (12.11.97)

(30) Priority Data:

60/029,444	12 November 1996 (12.11.96)	US
60/038,086	18 February 1997 (18.02.97)	US
60/059,368	19 September 1997 (19.09.97)	US
60/062,473	16 October 1997 (16.10.97)	US

(71) Applicant *(for all designated States except US)*: REGENTS OF THE UNIVERSITY OF MINNESOTA [US/US]; Morrill Hall, 100 Church Street S.E., Minneapolis, MN 55455 (US).

(71)(72) Applicants and Inventors: HOSTETTER, Margaret, K. [US/US]; 1832 Kenwood Parkway, Minneapolis, MN 55405 (US). CHENG, Qi [CN/US]; 11240 48th Avenue North, Plymouth, MN 55442 (US). FINKEL, David, J. [US/US]; Apartment 3, 414 S.E. Erie Street, Minneapolis, MN 55414 (US). SMITH, Beverly [US/US]; 1973 Sheridan Avenue South, Minneapolis, MN 55405 (US).

(74) Agent: MCCORMACK, Myra, H.; Mueting, Raasch, Gebhardt & Schwappach, P.A., P.O. Box 581415, Minneapolis, MN 55458–1415 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

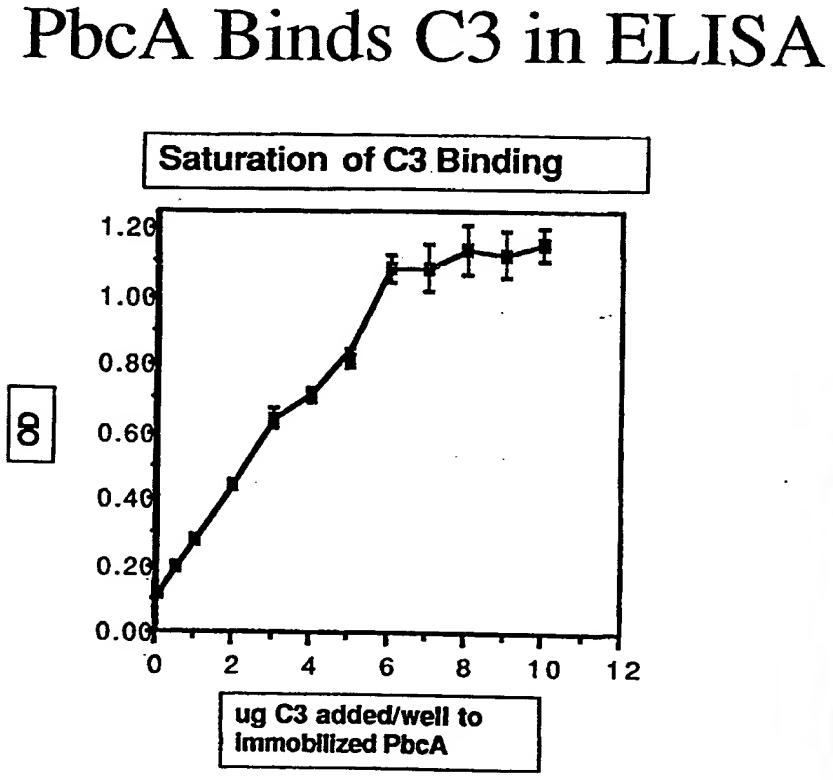
Published
With international search report.
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(88) Date of publication of the international search report:
23 July 1998 (23.07.98)

(54) Title: C3 BINDING PROTEIN OF *STREPTOCOCCUS PNEUMONIAE*

(57) Abstract

This invention relates to the identification of a human complement C3 binding protein from *Streptococcus pneumoniae* and to its sequence and to methods for its purification and use. The protein binds but does not degrade or cleave C3 and is implicated in *S. pneumoniae* virulence. The protein is recognized by antibodies produced by humans recovering from pneumococcal infection.



FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania
AM	Armenia
AT	Austria
AU	Australia
AZ	Azerbaijan
BA	Bosnia and Herzegovina
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CU	Cuba
CZ	Czech Republic
DE	Germany
DK	Denmark
EE	Estonia
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GH	Ghana
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IL	Israel
IS	Iceland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgyzstan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakstan
LC	Saint Lucia
LI	Liechtenstein
LK	Sri Lanka
LR	Liberia
LS	Lesotho
LT	Lithuania
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
MK	The former Yugoslav Republic of Macedonia
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
MX	Mexico
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SG	Singapore
SI	Slovenia
SK	Slovakia
SN	Senegal
SZ	Swaziland
TD	Chad
TG	Togo
TJ	Tajikistan
TM	Turkmenistan
TR	Turkey
TT	Trinidad and Tobago
UA	Ukraine
UG	Uganda
US	United States of America
UZ	Uzbekistan
VN	Viet Nam
YU	Yugoslavia
ZW	Zimbabwe

INTERNATIONAL SEARCH REPORT

Intern. ial Application No
PCT/US 97/20586

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/31 C12N1/21 C07K14/315 C07K16/12 A61K39/09
A61K39/40 //(C12N1/21,C12R1:46)

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C07K A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	WO 97 41151 A (UNIV ROCKEFELLER) 6 November 1997 see page 9, line 11 - line 19 see page 67, line 14 - line 25 see page 73, line 18 - page 74, line 8 SEQ ID no.24,25 see page 76, line 15 - page 78, line 13; claims 1-40	1-68
A	EP 0 622 081 A (UAB RESEARCH FOUNDATION) 2 November 1994 see the whole document	1-68
A	WO 93 24000 A (YOTHER JANET) 9 December 1993 see the whole document	1-68

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
8 May 1998	02.06.98

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Hornig, H

1

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/US 97/20586

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	ANONYMOUS: "A C3 - binding protein from Streptococcus pneumoniae." 97TH GENERAL MEETING OF THE AMERICAN SOCIETY FOR MICROBIOLOGY, MIAMI BEACH, FLORIDA, USA, MAY 4-8, 1997. ABSTRACTS OF THE GENERAL MEETING OF THE AMERICAN SOCIETY FOR MICROBIOLOGY 97 (0). 1997. 110. ISSN: 1060-2011, XP002064458 Accession no. B-478 ---	1-28, 33-47
P,X	C. ROSENOW ET AL.: "Contribution of novel choline-binding proteins to adherence, colonization and immunogenicity of Streptococcus pneumoniae" EMBL SEQUENCE DATABASE, 29 September 1997, HEIDELBERG, FRG, XP002064459 Accession no. AF019904; AC 030874; ---	1-28, 33-47
P,X	S. HAMMERSCHMIDT ET AL.: "SpsA, a novel pneumococcal surface protein with specific binding to immunglobulin A and secretory componnets" EMBL SEQUENCE DATABSE, 14 October 1997, HEIDELBERG, FRG, XP002064460 Accession no. AJ002054 -----	1-28, 33-47

1

INTERNATIONAL SEARCH REPORT

International application No.
PCT/US 97/20586

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

Although claims 55-63 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

2. [ ] Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

Information on patent family members

Inter nal Application No
PCT/US 97/20586

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9741151 A	06-11-1997	AU 2818297 A	19-11-1997
EP 0622081 A	02-11-1994	AU 682018 B	18-09-1997
		AU 5769694 A	27-10-1994
		CA 2116261 A	21-10-1994
		FI 941695 A	21-10-1994
		JP 7126291 A	16-05-1995
		NO 941420 A	21-10-1994
		US 5679768 A	21-10-1997
		ZA 9401584 A	12-10-1994
WO 9324000 A	09-12-1993	AU 4400293 A	30-12-1993

CORRECTED VERSION*

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 :
C12N 15/31, 1/21, C07K 14/315, 16/12, A61K 39/09, 39/40 // (C12N 1/21, C12R 1:46)

A3

(11) International Publication Number: WO 98/21337

(43) International Publication Date: 22 May 1998 (22.05.98)

(21) International Application Number: PCT/US97/20586

(22) International Filing Date: 12 November 1997 (12.11.97)

(30) Priority Data:

60/029,444	12 November 1996 (12.11.96)	US
60/038,086	18 February 1997 (18.02.97)	US
60/059,368	19 September 1997 (19.09.97)	US
60/062,473	16 October 1997 (16.10.97)	US

(71) Applicant *(for all designated States except US)*: REGENTS OF THE UNIVERSITY OF MINNESOTA [US/US]; Morrill Hall, 100 Church Street S.E., Minneapolis, MN 55455 (US).

(71)(72) Applicants and Inventors: HOSTETTER, Margaret, K. [US/US]; 1832 Kenwood Parkway, Minneapolis, MN 55405 (US). CHENG, Qi [CN/US]; 11240 48th Avenue North, Plymouth, MN 55442 (US). FINKEL, David, J. [US/US]; Apartment 3, 414 S.E. Erie Street, Minneapolis, MN 55414 (US). SMITH, Beverly [US/US]; 1973 Sheridan Avenue South, Minneapolis, MN 55405 (US).

(74) Agent: MCCORMACK, Myra, H.; Mueting, Raasch, Gebhardt & Schwappach, P.A., P.O. Box 581415, Minneapolis, MN 55458–1415 (US).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, ML, MR, NE, SN, TD, TG).

Published

With international search report.

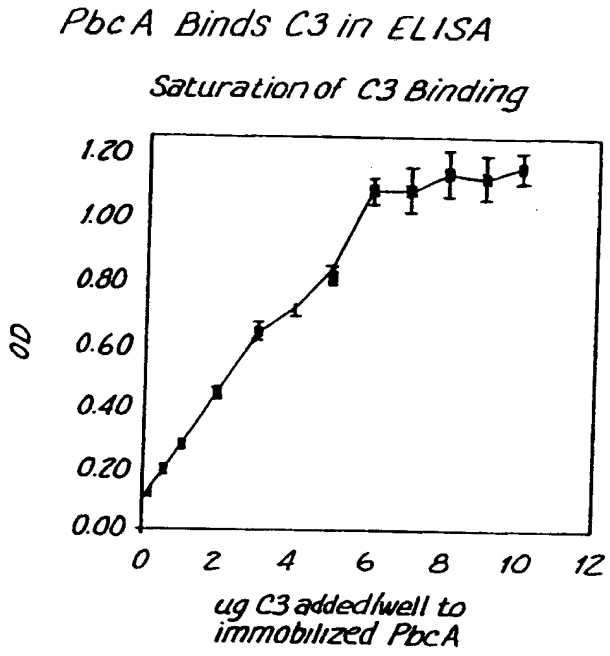
Before the expiration of the time limit for amending the claims and to be republished in the event of the receipt of amendments.

(88) Date of publication of the international search report:
23 July 1998 (23.07.98)

(54) Title: C3 BINDING PROTEIN OF *STREPTOCOCCUS PNEUMONIAE*

(57) Abstract

This invention relates to the identification of a human complement C3 binding protein from *Streptococcus pneumoniae* and to its sequence and to methods for its purification and use. The protein binds but does not degrade or cleave C3 and is implicated in *S. pneumoniae* virulence. The protein is recognized by antibodies produced by humans recovering from pneumococcal infection.



*(Referred to in PCT Gazette No. 34/1998, Section II)

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	**ES**	Spain	**LS**	Lesotho	**SI**	Slovenia
AM	Armenia	**FI**	Finland	**LT**	Lithuania	**SK**	Slovakia
AT	Austria	**FR**	France	**LU**	Luxembourg	**SN**	Senegal
AU	Australia	**GA**	Gabon	**LV**	Latvia	**SZ**	Swaziland
AZ	Azerbaijan	**GB**	United Kingdom	**MC**	Monaco	**TD**	Chad
BA	Bosnia and Herzegovina	**GE**	Georgia	**MD**	Republic of Moldova	**TG**	Togo
BB	Barbados	**GH**	Ghana	**MG**	Madagascar	**TJ**	Tajikistan
BE	Belgium	**GN**	Guinea	**MK**	The former Yugoslav Republic of Macedonia	**TM**	Turkmenistan
BF	Burkina Faso	**GR**	Greece			**TR**	Turkey
BG	Bulgaria	**HU**	Hungary	**ML**	Mali	**TT**	Trinidad and Tobago
BJ	Benin	**IE**	Ireland	**MN**	Mongolia	**UA**	Ukraine
BR	Brazil	**IL**	Israel	**MR**	Mauritania	**UG**	Uganda
BY	Belarus	**IS**	Iceland	**MW**	Malawi	**US**	United States of America
CA	Canada	**IT**	Italy	**MX**	Mexico	**UZ**	Uzbekistan
CF	Central African Republic	**JP**	Japan	**NE**	Niger	**VN**	Viet Nam
CG	Congo	**KE**	Kenya	**NL**	Netherlands	**YU**	Yugoslavia
CH	Switzerland	**KG**	Kyrgyzstan	**NO**	Norway	**ZW**	Zimbabwe
CI	Côte d'Ivoire	**KP**	Democratic People's Republic of Korea	**NZ**	New Zealand		
CM	Cameroon			**PL**	Poland		
CN	China	**KR**	Republic of Korea	**PT**	Portugal		
CU	Cuba	**KZ**	Kazakstan	**RO**	Romania		
CZ	Czech Republic	**LC**	Saint Lucia	**RU**	Russian Federation		
DE	Germany	**LI**	Liechtenstein	**SD**	Sudan		
DK	Denmark	**LK**	Sri Lanka	**SE**	Sweden		
EE	Estonia	**LR**	Liberia	**SG**	Singapore		

C3 BINDING PROTEIN OF *STREPTOCOCCUS PNEUMONIAE*

Field of the Invention

This invention relates to *Streptococcus pneumoniae* and in particular this invention relates to the identification of an *S. pneumoniae* protein that is implicated in *S. pneumoniae* virulence and is capable of binding the complement protein, C3.

Background of the Invention

This application is a continuation in part of a provisional application filed October 16, 1997 (U.S. Serial No. 60/062,473) entitled "C3 BINDING PROTEIN OF *S. PNEUMONIAE*." This application is also a continuation in part of a provisional application filed on September 19, 1997 (U.S. Serial No. 60/059,368), a provisional application filed on February 18, 1997 (U.S. Serial No. 60/038,086) and a provisional application filed on November 12, 1996 (U.S. Serial No. 60/029,444) all entitled "C3 BINDING PROTEIN OF *S. PNEUMONIAE*."

Respiratory infection with the bacterium *Streptococcus pneumoniae* (*S. pneumoniae*) leads to an estimated 500,000 cases of pneumonia and 47,000 deaths annually. Those persons at highest risk of bacteremic pneumococcal infection are infants under two years of age and the elderly. In these populations, *S. pneumoniae* is the leading cause of bacterial pneumonia and meningitis. Moreover, *S. pneumoniae* is the major bacterial cause of ear infections in children of all ages. Both children and the elderly share defects in the synthesis of protective antibodies to pneumococcal capsular polysaccharide after either bacterial colonization, local or systemic infection, or vaccination with purified polysaccharides. *S. pneumoniae* is the leading cause of invasive bacterial respiratory disease in both adults and children with HIV infection and produces hematogenous infection in these patients (Connor et al. *Current Topics in AIDS* 1987;1:185-209 and Janoff et al. *Ann. Intern. Med.* 1992;117(4):314-324).

Individuals who demonstrate the greatest risk for severe infection are not able to make antibodies to the current capsular polysaccharide vaccines. As a result, there are now four conjugate vaccines in clinical trial. Conjugate vaccines consist of pneumococcal capsular polysaccharides coupled to protein carriers or adjuvants in an attempt to boost the antibody response. However, there are other potential problems with conjugate vaccines currently in clinical trials. For example, pneumococcal serotypes that are most prevalent in the United States are different from the serotypes that are most common in places such as Israel, Western Europe, or Scandinavia. Therefore, vaccines that may be useful in one geographic locale may not be useful in another. The potential need to modify currently available capsular polysaccharide vaccines or to develop protein conjugates for capsular vaccines to suit geographic serotype variability entails prohibitive financial and technical complications. Thus, the search for immunogenic, surface-exposed proteins that are conserved worldwide among a variety of virulent serotypes is of prime importance to the prevention of pneumococcal infection and to the formulation of broadly protective pneumococcal vaccines. Moreover, the emergence of penicillin and cephalosporin-resistant pneumococci on a worldwide basis makes the need for effective vaccines even more exigent (Baquero et al. *J. Antimicrob. Chemother.* 1991;28S;31-8).

Several pneumococcal proteins have been proposed for conjugation to pneumococcal capsular polysaccharide or as single immunogens to stimulate immunity against *S. pneumoniae*. Surface proteins that are reported to be involved in adhesion of *S. pneumoniae* to epithelial cells of the respiratory tract include PsaA, PspC/CBP112, and IgA1 proteinase (Sampson et al. *Infect. Immun.* 1994;62:319-324, Sheffield et al. *Microb. Pathogen.* 1992; 13: 261-9, and Wani, et al. *Infect. Immun.* 1996; 64:3967-3974). Antibodies to these adhesins could inhibit binding of pneumococci to respiratory epithelial cells and thereby reduce colonization. Other cytosolic pneumococcal proteins such as pneumolysin, autolysin, neuraminidase, or hyaluronidase are proposed as vaccine antigens because antibodies could potentially block the toxic effects of

these proteins in patients infected with *S. pneumoniae*. However, these proteins are typically not located on the surface of *S. pneumoniae*, rather they are secreted or released from the bacterium as the cells lyse and die (Lee et al. *Vaccine* 1994; 12:875-8 and Berry et al. *Infect. Immun.* 1994; 62:1101-1108). While use of these cytosolic proteins as immunogens might ameliorate late consequences of *S. pneumoniae* infection, antibodies to these proteins would neither promote pneumococcal death nor prevent pneumococcal colonization.

A prototypic surface protein that is being tested as a pneumococcal vaccine is the pneumococcal surface protein A (PspA). PspA is a heterogeneous protein of about 70-140 kDa. The PspA structure includes an alpha helix at the amino terminus, a proline-rich sequence in the mid-portion of the protein, and terminates in a series of choline-binding repeats at the carboxy-terminus. Although much information regarding its structure is available, PspA is not structurally conserved among a variety of pneumococcal serotypes, and its function is entirely unknown (Yother et al. *J. Bacteriol.* 1992;174:601-9 and Yother *J. Bacteriol.* 1994;176:2976-2985). Studies have confirmed the immunogenicity of PspA in animals (McDaniel et al. *Microb. Pathogen.* 1994; 17:323-337). Despite the immunogenicity of PspA, the heterogeneity of PspA, its existence in four structural groups (or clades), and its uncharacterized function complicate its ability to be used as a vaccine antigen.

In patients who cannot make protective antibodies to the type-specific polysaccharide capsule, the third component of complement, C3, and the associated proteins of the alternative complement pathway constitute the first line of host defense against *S. pneumoniae* infection. Because complement proteins cannot penetrate the rigid cell wall of *S. pneumoniae*, deposition of opsonic C3b on the pneumococcal surface is the principal mediator of pneumococcal clearance. Interactions of pneumococci with plasma C3 are known to occur during pneumococcal bacteremia, when the covalent binding of C3b, the opsonically active fragment of C3, initiates phagocytic recognition and ingestion (Johnston et al. *J. Exp. Med* 1969;129:1275-1290, Hasin HE, *J. Immunol.* 1972; 109:26-31 and Hostetter et al. *J. Infect. Dis.* 1984; 150:653-61).

C3b deposits on the pneumococcal capsule, as well as on the cell wall. This method for controlling *S. pneumoniae* infection is fairly inefficient and could be beneficially amplified by the presence of antibodies to surface components of *S. pneumoniae*. There currently exists a strong need for methods and therapies to limit *S. pneumoniae* infection.

Summary of the Invention

The present invention relates to the identification and purification of an about 90 kDa to about 110 kDa (± 5 kDa) protein, as determined following electrophoresis on a 15% SDS-PAGE gel. The protein is named PbcA and is isolatable from *S. pneumoniae* strains that are capable of binding to human complement protein C3. The protein, PbcA, comprises an amino terminus containing region comprising SEQ ID NO:1 and is capable of binding but not cleaving or degrading the human complement protein C3. The protein also comprises a proline rich region and in one embodiment is a surface exposed protein of *S. pneumoniae*.

This invention also relates to the production of antibodies specifically recognizing PbcA. In one embodiment the antibodies are polyclonal and in another embodiment the antibodies are monoclonal. The antibodies can be produced by immunizing a mammal with all or a portion of PbcA. In one embodiment, the monoclonal antibodies are rodent derived.

In another aspect of this invention a method is provided for generating an immune response to *S. pneumoniae in vivo* comprising the steps of: administering a protein or an immunogenic fragment of a protein from *S. pneumoniae* to an animal wherein the amino terminus containing region of the protein comprises SEQ ID NO:1. In one embodiment, the protein is capable of binding but not cleaving or degrading the human complement protein C3. Preferably, the method further comprises detecting an immune response to *S. pneumoniae* in the mammal. Preferably, the immune response comprises the production of antibodies to *S. pneumoniae*. The animal can be a mouse, rat, chinchilla, a rabbit or a human and the method can further comprise the steps of

isolating antibody producing cells from the mammal and preparing monoclonal antibodies to the C3 binding protein.

In yet another aspect of the invention a method is disclosed for obtaining a purified C3 binding protein from *S. pneumoniae* comprising the steps of: obtaining a protein sample from *S. pneumoniae*; precipitating the protein to form a precipitate; applying the precipitate to a Thiopropyl Sepharose 6B affinity chromatography column comprising methylamine-treated C3; and eluting the C3 binding protein from the column using an elution buffer comprising about 20% ethanol. In one aspect of this embodiment the invention relates to C3 binding protein preparable by these methods.

This invention also relates to a C3 binding protein having the sequence of SEQ ID NO:6 and to isolated nucleic acid encoding a C3 binding protein and comprising the DNA sequence of SEQ ID NO:5 and to isolated nucleic acid having the DNA sequence of SEQ ID NO:4.

In another aspect of this invention, the invention relates to isolated nucleic acid encoding C3 binding protein of about 90 kDa to about 110 kDa (± 5 kDa), in one embodiment, and comprising nucleic acids 1-1500 of SEQ ID NO:5 and to an isolated nucleic acid sequence encoding the C3 binding protein wherein the protein exhibits C3 binding activity and wherein the protein comprises at least 80% nucleic acid homology to nucleic acids 1-1500 of SEQ ID NO:5. Preferably, the nucleic acid homology is at least 95%.

The invention also relates to a C3 binding protein isolatable from *S. pneumoniae* having DNA that is hybridizable to a nucleic acid fragment of at least 500 bp from nucleic acids 1-1500 of SEQ ID NO: 5 under hybridization conditions of about 6X SSC, 5X Denhardt's, 0.5% SDS, 100 µg/ml denatured, fragment salmon sperm DNA overnight at 65°C and washed in 2X SSC, 0.1%SDS, one time at room temperature for about 10 mn, followed by one time at 65°C for about 15 mn and followed by at least one wash in 0.2 X SSC, 0.1% SDS at room temperature for about 3-5 minutes. Preferably this protein further comprises at least 2 choline binding repeat and still more preferably the protein further comprises at least 2 choline binding repeats.

The invention also relates to peptide fragments of at least 15 bp from SEQ ID NO:5 and to insertion and deletion mutants that do not express PbcA.

In another aspect of this invention, the invention relates to PbcA proteins. In one embodiment the invention relates to an isolated protein comprising SEQ ID NO:1 and at least two choline binding repeats. Preferably the protein is isolated from *S. pneumoniae* and also preferably the protein binds human complement protein C3. In one version of this embodiment, the protein is a recombinant protein or a purified protein from *S. pneumoniae*. Preferably the protein has a molecular weight as determined on a 15% polyacrylamide gel of between about 90 kDa to about 110 kDa and preferably the protein comprises a proline rich region. Additionally the protein can comprise SEQ ID NO:2 or SEQ ID NO:3.

Alternatively, the protein of this invention can comprise SEQ ID NO:6 or the protein can be an isolated protein capable of binding to, but not cleaving or degrading, human complement C3 and wherein the protein comprises SEQ ID NO:1. Preferably the protein is isolated from *S. pneumoniae* and in one embodiment, the protein further comprises a proline rich region. The protein can further comprise SEQ ID NO:2. Preferably the protein has at least about 95% homology to a C3 binding protein from *S. pneumoniae* and also preferably, the protein has a molecular weight as determined on a 15% polyacrylamide gel of between about 90 kDa to about 110 kDa. In one aspect, the isolated protein is a recombinant protein and in another, the protein is isolated from *S. pneumoniae* bacteria.

In another embodiment of the proteins of this invention, the invention relates to a recombinant protein comprising SEQ ID NO:1, wherein the protein has a molecular weight as determined on a 15% polyacrylamide gel of between about 90 kDa to about 110 kDa. Preferably the protein binds human complement protein C3. The protein can further include a proline rich region and preferably the protein does not cleave or degrade human complement protein C3.

In another embodiment of the proteins of this invention, the protein comprises amino acids 1-410 of SEQ ID NO:6.

In yet another embodiment of the proteins of this invention, the invention relates to a protein that binds, but does not cleave or degrade, human complement protein C3, wherein nucleic acid encoding the protein hybridizes to SEQ ID NO:4 under hybridization conditions of 6X SSC, 5X Denhardt, 0.5% SDS, and 100 μg/ml fragmented and denatured salmon sperm DNA hybridized overnight at 65°C and washed in 2X SSC, 0.1% SDS one time at room temperature for about 10 minutes followed by one time at, 65°C for about 15 minutes followed by at least one wash in 0.2XSSC, 0.1% SDS at room temperature for at least 3-5 minutes. Preferably, the protein further comprises SEQ ID NO:1 and optionally, the protein can comprise SEQ ID NO:2 or SEQ ID NO:3. Preferably the protein is at least 15 amino acids in length. Preferably the protein comprises a proline rich region. In one aspect of this embodiment, the protein is a recombinant protein and in another, the protein is a synthetic peptide. In one aspect of this embodiment, the protein is a peptide of at least 15 amino acids from SEQ ID NO:6. The proteins can also be used to create antibody and the proteins of this embodiment can be used to generate antibody capable of specifically binding to the protein. In one embodiment, the antibody is a monoclonal antibody and in another the antibody is a polyclonal antibody. Preferably the monoclonal antibody is at least partially rodent-derived.

This invention also relates to nucleic acid encoding the proteins of this invention. In one embodiment, the nucleic acid of this invention encodes a protein comprising at least two choline binding domains and SEQ ID NO: 1. Preferably the protein encoded by the nucleic acid further comprising a proline rich region. Also preferably, the nucleic acid is isolated from an *S. pneumoniae* genome. Preferably, the nucleic acid is capable of hybridizing to SEQ ID NO:4 and in another embodiment, the protein encoded by the nucleic acid binds to human complement protein C3. In one aspect of this embodiment, the nucleic acid is positioned in a nucleic acid vector. Preferably the vector is an expression

vector and the expression vector directs expression of the protein by the nucleic acid.

In another embodiment of the nucleic acid of this invention, the invention relates to isolated nucleic acid encoding a protein comprising SEQ ID NO:1 and a proline rich region wherein the protein encoded by the nucleic acid binds but does not cleave or degrade human complement C3.

In yet another embodiment, the invention relates to isolated nucleic acid fragment encoding an about 90 kDa to about 110 kDa protein with C3 binding activity, wherein the nucleic acid fragment has at least 80% homology to at least 500 bp from nucleic acids 1-1500 of SEQ ID NO:5 and in another embodiment, the isolated nucleic acid fragment comprises base pairs 1-1500 of SEQ ID NO:5.

The invention also relates to a method for isolating a C3 binding protein from a bacterium comprising the steps of: obtaining a protein sample from a bacterium; applying the sample to a solid support comprising methylamine treated complement protein C3; washing the solid support; and removing a C3 binding protein from the solid support in a solution comprising alcohol wherein the C3 binding protein does not cleave or degrade C3. Preferably the bacterium is *S. pneumoniae* and in another embodiment, the bacterium is *E. coli*. Preferably the solid support comprises an affinity column and preferably the alcohol is ethanol. In one embodiment, the solution comprising alcohol is a buffer comprising 20% ethanol. The invention also relates to C3 binding protein preparable by this method.

The invention also relates to a method for producing an immune response to *S. pneumoniae* comprising the steps of: administering a therapeutically effective amount of at least a portion of a protein to a mammal, wherein the protein binds but does not cleave or degrade human complement protein C3 and, wherein nucleic acid encoding the protein hybridizes to SEQ ID NO:4 under hybridization conditions of 6XSSC, 5X Denhardt, 0.5% SDS, and 100 μg/ml fragmented and denatured salmon sperm DNA, hybridized overnight at 65°C and washed in 2x SSC, 0.1% SDS one time at room temperature for

about 10 minutes followed by one time at 65°C for about 15 minutes followed by at least one wash in 0.2xSSC, 0.1% SDS at room temperature for at least 3-5 minutes; and detecting an immune response to the protein. In one embodiment, the protein is at least 15 amino acids in length and in another the protein is a chimeric protein. In yet another embodiment, the protein comprises SEQ ID NO:1. In another embodiment, the protein has a molecular weight on a 15% polyacrylamide gel of between about 90 kDa to about 110 kDa.

The invention also relates to a method for reducing *S. pneumoniae* binding to C3 comprising the steps of: administering a therapeutically effective amount of at least a portion of an antibody to a mammal, wherein the antibody specifically recognizes a C3 binding protein from *S. pneumoniae*, and wherein the C3 binding protein is a protein that binds, but does not cleave or degrade human complement protein C3. In one embodiment the antibody comprises at least one variable domain from a monoclonal antibody. In another embodiment, the antibody is administered to the air passages of the mammal or intravenously.

In yet another aspect of this invention, the invention relates to a non-naturally occurring *S. pneumoniae* bacterium that does not express a detectable human complement C3 binding protein, wherein the C3 binding protein comprises SEQ ID NO:1. In one embodiment, the bacterium is the product of an insertion into the gene encoding the C3 binding protein and in another, the bacterium is a product of a deletion in the gene encoding the C3 binding protein. In yet another embodiment, the bacterium is produced by homologous recombination of the gene encoding the C3 binding protein gene with at least a portion of a non-native C3 binding protein gene and in one aspect of this embodiment, the gene encoding at least a portion of the non-native C3 binding protein comprises a mutation within the non-native gene.

Brief Description of the Figures

Fig. 1 is an amino acid alignment of the choline binding regions of PbcA from *S. pneumoniae* strains CP1200, R6x and 23F (SEQ ID NOS:9, 10, and11).

Fig. 2 is a graph illustrating PbcA binding to human C3 by ELISA.

Fig. 3 summarizes studies to assess the ability of antibody to PbcA to inhibit C3 binding. Figure 3a is a Western blot illustrating the specificity of the antibody preparation for PbcA from CP1200. Figure 3b is a graph illustrating the ability of antibody to PbcA to inhibit C3 binding by ELISA.

Fig. 4 illustrates an exemplary strategy for producing a *pbc*A insertion mutant.

Fig. 5a is a photograph of a Southern blot of DNA from transformed and nontransformed *S. pneumoniae* cells probed with a 1.5kb fragment of *pbc*A. Figure 5b is a photograph of a Southern blot probed with vector pVA891.

Figure 6 is a photograph of Western blot experiments demonstrating that the insertion-duplication mutants of this invention do not produce PbcA protein capable of binding C3 or antibody to PbcA.

Detailed Description of the Preferred Embodiments

The present invention provides a *S. pneumoniae* protein that binds C3. The protein is about 90 kDa to about 110 kDa (± about 5 kDa) in mass as observed following electrophoresis on a 15% SDS-PAGE gel and is referred to in this disclosure as "PbcA." PbcA was initially identified using the methods of Example 1. Supernatant proteins from exponentially growing cultures of *S. pneumoniae* were tested by Western blot for their ability to bind to purified human C3 following electrophoresis on a 15% SDS-PAGE gel. In the *S. pneumoniae* strain CP1200, both lysates and supernatants from exponentially growing cells produced a protein band, identified as PbcA, at approximately 90 kDa on 15% SDS-PAGE gels under non-reducing conditions. The protein bound methylamine-treated human C3 (produced as described by Hostetter MK, et al. *J. Inf. Dis.* 150:653-661, 1984), labeled with biotin. PbcA is present in lysates and supernatants from a variety of other *S. pneumoniae* strains. The size of PbcA

varies on a 15% SDS-PAge gel in *S. pneumoniae* strains from about 90 kDa to about 110 kDa. All PbcA proteins identified thus far are capable of binding, but not cleaving or degrading, purified human complement protein C3. In some strains protein bands of smaller molecular weight are observed.

As will be recognized by those of skill in the art, there are a variety of methods for isolating an individual protein from bacterial supernatants. As one example, the *S. pneumoniae* C3 binding protein, PbcA, is isolated using the methods of Example 2. PbcA, can be purified from other secreted pneumococcal proteins by affinity chromatography. Secreted proteins can be precipitated in a final concentration of 10% trichloroacetic acid (TCA) at 4°C overnight according to Example 2. Resuspended proteins from the TCA precipitate are subjected to affinity column chromatography using methylamine-treated human C3 (*supra*). Elution of the PbcA protein from the affinity column has proven to be difficult. Surprisingly, the PbcA protein from the column can be eluted using an elution buffer comprising an alcohol, preferably ethanol and more preferably about 20% ethanol in the Tris-HCl/NaCl wash buffer. In view of this disclosure, now that PbcA has been identified, those skilled in the art will recognize that other methods could be used to identify, isolate and purify the protein from a variety of C3-binding *S. pneumoniae* without undue experimentation.

Multiple eluates can be pooled to obtain sufficient sample for further analysis. As one example, a sample can be electrophoresed on an SDS-PAGE gel and transferred to nitrocellulose. The protein can be subjected to amino terminal analysis and tryptic digestion for internal peptide sequencing. The following sequences were obtained from the tryptic digest analysis:

A peptide positioned near the amino terminus:

TENEGSTQAATSSNMAKTEH (SEQ ID NO:1)

And internal regions:

EKPAEQPQPAPATQP (SEQ ID NO:2)

SSDSSVGEETLPSSSLK (SEQ ID NO:3)

SEQ ID NO:2 is proline rich and has at least a 75% homology with the *S. pneumoniae* protein PspA over 13 amino acids. Although the proline rich region of PspA aligned with SEQ ID NO:2, neither SEQ ID NO:1 nor SEQ ID NO:3 had any substantial homology to any proteins or peptides previously published in the GenBank database (less than 35% homology). The term "proline rich" as used herein refers to a protein having a stretch of amino acids having at least 5 proline amino acids over a total of about 15 amino acids.

These sequences were confirmed in *S. pneumoniae* strain CP1200 following isolation of the gene and sequencing to obtain the nucleic acid sequence encoding PbcA. In one embodiment of this invention, PbcA is an about 90 kDa to about 110 kDa (± 5 kDa, meaning about 85 kDa to about 115 kDa) when *S. pneumoniae* proteins are separated on a 15% SDS-PAGE gel, and in another embodiment, the protein further includes SEQ ID NO:2 and SEQ ID NO:3.

Oligonucleotides corresponding in whole or in part to SEQ ID NOS: 1-3 are useful for identifying and isolating the nucleic acid encoding PbcA (the gene encoding PbcA is termed *pbcA*) and for isolating the *pbc*A and PbcA from a variety of *S. pneumoniae* strains. For example, oligonucleotides corresponding in whole or in part from SEQ ID NOS 1-3 can be used to amplify sequences from genomic DNA isolated from *S. pneumoniae* using standard polymerase chain reaction technology. The amplified sequences can then be directly used as probes or the amplified sequences can be incorporated into a vector for plasmid amplification in a suitable host such as a bacteria or virus and then isolated for sequencing, cloning and for use as probes to detect DNA from libraries of *S. pneumoniae*. The DNA isolated from these procedures is useful in sequencing reactions to obtain the nucleic acid sequence encoding PbcA and to produce vectors, such as expression vectors encoding PbcA as well as for producing recombinant protein. Example 5 provides a preferred method for isolating nucleic acid encoding PbcA.

PbcA can be expressed as a recombinant protein or isolated from S. pneumoniae lysates. The *S. pneumoniae* C3 binding protein, PbcA, binds C3

without cleaving or degrading the C3 molecule. A number of bacterial proteins have been reported to bind and to cleave C3 or other complement proteins. For example, a 140 kDa C5a peptidase from group A streptococci cleaves a His_{67}-Lys_{68} bond at the carboxy terminus of C5a, thereby abolishing the chemoattractant capabilities of the molecule. An enzyme related to the C5a peptidase is also found in group B streptococci (Cleary et al. *Infect. Immun.* 1992; 60:4239-4244 and Bohnsack et al. *Biochim. et Biophys. Acta* 1991; 1079:222-228). Production of an elastase-like enzyme, as can be seen with 24-hour culture supernatants from *Pseudomonas aeruginosa* (Suter et al. *J. Infect. Dis.* 1984; 149:523-31), cleaves C3 into characteristic fragments of 66 kDa and 100 kDa from the C3 α-chain. Like the elastase-like enzyme from *P. aeruginosa*, a 56 kDa neutral cysteine proteinase from *Entamoeba histolytica* cleaves C3a between residues Ser_{78}/Asn_{79}, yielding a defined C3 cleavage fragment of 105 kDa (Reed, et al. *J. Immunol.* 1989; 143:189-95).

C3-cleaving proteinases have been isolated from the membranes of some mammalian cells, including human erythrocytes (p57), neutrophils, and melanoma cells resistant to complement-mediated killing. These proteins are typically serine proteases which yield defined cleavage fragments. For example, p57 cleaves both the α- and β-chains of C3, while the melanoma proteinase cleaves only the α'-chain of C3b, generating a fragment of 35 kDa (Charriaut-Marlangue et al. *Biochem. Biophys. Res. Commun.* 1986; 140:1113-1120 and Ollert et al. *J. Immunol.* 1990; 144:3862-7).

In contrast to these studies, C3 cleavage and/or degradation was not observed with PbcA. PbcA binds C3 without the production of defined cleavage fragments and without evidence of degradation of C3. Although there are microbial precedents for binding and cleaving of complement proteins, there is no previously reported microbial protein that binds C3 non-covalently without degrading or cleaving the molecule. The interaction of C3 and other complement components with proteins from group A and B streptococci, *P. aeruginosa*, and amoebae appears to be quite distinct from what has been observed with *S. pneumoniae*.

Table 1 (SEQ ID NO:4) provides a nucleic acid sequence encompassing the open reading frame encoding a 90 kDa PbcA protein from *S. pneumoniae* strain CP1200. Untranslated 5' and 3' regions are also included in SEQ ID NO:4. The open reading frame encoding PbcA begins at nucleotide 383 and ends with nucleotide 2074. Table 2 is a map providing the amino acids (SEQ ID NO: 6) encoded by the nucleic acids of the major open reading frame from SEQ ID NO:4 (provided in Table 2 as SEQ ID NO:5). The protein predicted from SEQ ID NO:4 contains an amino terminus containing segment with C3 binding activity (upstream from the choline binding repeat region) and a series of choline binding repeats (beginning in SEQ ID NO:5 at about nucleic acid position 1501).

Primers were selected from SEQ ID NO:4 to span the choline binding repeats to assess choline binding repeat variability between strains. The primers used were:

5' GCACAACCATCTACTCCA 3' (SEQ ID NO: 7), and

5' GTACAGGAATTCAGTATTAACTA 3' (SEQ ID NO:8)

Amplification reactions were performed using DNA from three different *S. pneumoniae* strains: CP1200, R6x and virulent strain 23F (obtained from Dr. Steve Pelton, Boston City Hospital, Boston, MA and identified as isolate "freezer #365"). The results of the amplification studies indicated that the number of choline binding repeats varied depending on the *S. pneumoniae* strain. For example, strain CP1200 contained about 4 repeats while 23F contained about 8 repeats and R6X contained at least about 10 choline binding repeats using the Yother et al. model for choline binding repeat regions (see *infra*). All *S. pneumoniae* strains studied thus far have at least two choline binding repeats. Therefore, in another embodiment of this invention, PbcA is a C3-binding protein from *S. pneumoniae* including SEQ ID NO:1 and at least 2 choline binding repeats.

Choline binding domains are known in the art and a number of references discuss the characteristics of a variety choline binding domains (see, for example, Du Clos et al. *J. Biol. Chem.* 266(4):2167-2171, 1991; Liu et al. *J.*

Biol. Chem. 266(22):14813-14821, 1991; Agrawal et al. *J. Biol.Chem.* 267(35):25352-25358, 1992; and *Nature Structural Biology* 3(4):346-354, 1994). A choline-binding repeat sequence has been identified in *S. pneumoniae* protein PspA as TGWKQENGMWYFYNTDGSMA (SEQ ID NO:12) (see Yother, J. and White, JM, *J. Bacteriology* 176:2976-2985, 1994) Choline binding repeats are associated with membrane binding proteins. Without intending to limit this invention, the strains studied thus far indicate that the virulent strains appear to have more choline binding repeats on average than nonvirulent strains. At the very least it appears that *S. pneumoniae* strains show considerable variability in the choline binding repeat region.

These results are consistent with electrophoretic studies assessing variability in the size of PbcA proteins obtained from a number of *S. pneumoniae* strains. Proteins isolated according to the methods of this invention, when separated by SDS-PAGE, demonstrate some size variability (See Fig. 6) from about 90 kDa to about 110 kDa (± 5 kDa). This variability can be attributed, at least in part, to the variability in length of the choline binding repeat region. The proteins of this invention also include a series of peptides at the carboxy terminus of the choline binding repeat region.

Fig. 1 provides a comparison of the amino acid sequence from a region containing the choline binding repeat region from PbcA for three different *S. pneumoniae* strains: CP1200, R6x and 23F(BD23). The amino acids in the boxes represent variations in the 23F (virulent strain BD23) choline binding domain repeat region (SEQ ID NOS: 9, 10, and 11).

One example of a fragment containing nucleic acid encoding PbcA protein is provided in SEQ ID NO:4. The nucleic acid sequence encoding PbcA in strain CPI200 is provided as SEQ ID NO:5 (see Table 2). In another aspect of this invention, a protein of this invention has the amino acid sequence of SEQ ID NO:6. In one embodiment, a protein of this invention includes amino acids 1-410 of SEQ ID NO:6. In addition, variability between strains has been identified at the amino acid and nucleic acid level. For example, the amino terminus contains some variability and SEQ ID NO:1 may lack the first

threonine residue. In general, PbcA proteins include C3 binding proteins from *S. pneumoniae* that preferably have at least 80% nucleic acid homology within the DNA of the amino terminus-containing region (the region amino-terminal to the choline binding repeats) to SEQ ID NO:5. More preferably, the PbcA proteins have at least 95% homology to the amino terminus-containing region of PbcA and still more preferably the PbcA protein includes SEQ ID NO:1. In another embodiment the PbcA proteins additionally include SEQ ID NO:2 or SEQ ID NO:3.

In one example, the nucleic acid encoding PbcA can be obtained from a variety of *S. pneumoniae* strains. A *S. pneumoniae* genomic library can be prepared using *S. pneumoniae* genomic DNA and in a preferred example, an *S. pneumoniae* genomic library was prepared using the CP1200 strain. Custom libraries can be obtained using a variety of standard methods for library construction. In these studies genomic DNA from *S. pneumoniae* was given to a commercial custom library supplier (Stratagene, LaJolla, CA). In one example, *S. pneumoniae* CP1200 strain genomic DNA was used to prepare the library (Example 5) The results of these studies identified a cloned nucleic acid fragment encoding PbcA. This fragment is useful for identifying PbcA encoding nucleic acid in other *S. pneumoniae* strains and the nucleic acid can be incorporated into vectors including expression vectors, for example to produce recombinant protein using methods such as those described by Sambrook et al. (cited below).

PbcA is preferably encoded by nucleic acid that is capable of hybridizing to at least 500 bp from the amino terminus region of SEQ ID NO: 5 under hybridization conditions of about 6X SSC, 5X Denhardt's, 0.5% SDS, 100 µg/ml denatured, fragmented salmon sperm DNA hybridized overnight at 65°C and washed in 2X SSC, 0.1%SDS, one time at room temperature for about 10 mn, followed by one time at 65°C for about 15 mn and followed by at least one wash in 0.2 X SSC, 0.1% SDS at room temperature for about 3-5 minutes. The protein of this invention preferably includes at least two and preferably at least three choline binding repeats.

This invention also relates to nucleic acid fragments of at least 20 bp from SEQ ID NO:5 and to isolated nucleic acid fragments of at least 100 base pairs that hybridize to the *S. pneumoniae* genome or to SEQ ID NO:5 under the hybridization conditions of about 6X SSC, 5X Denhardt's, 0.5% SDS, 100 µg/ml denatured, fragment salmon sperm DNA hybridized overnight at 65°C and washed in 2X SSC, 0.1%SDS, one time at room temperature for about 10 min., followed by one wash at 65°C for about 15 min and followed by at least one wash in 0.2 X SSC, 0.1% SDS at room temperature for about 3-5 minutes. Preferably the nucleic acid fragments of at least 100 base pairs are part of a nucleic acid sequence that encodes a C3 binding protein that is at about 90 kDa to about 110 kDa. Preferably the nucleic acid fragment encodes a protein having at least two choline binding repeats.

As demonstrated in Example 3, the immunogenicity of PbcA can be studied by testing for the presence of antibodies to PbcA in acute and convalescent sera from patients with culture-proven pneumococcal infection. *S. pneumoniae* proteins can be separated by SDS-PAGE electrophoresis, transferred to nitrocellulose, and incubated with acute or convalescent serum (standard Western blot procedure). Results to detect the presence of antibodies to PbcA in human sera indicated that acute serum did not contain antibodies to any pneumococcal proteins released into the supernatant but that convalescent serum contained antibodies that recognized a band on an acrylamide gel with a molecular weight of PbcA.

PbcA, or fragments of PbcA, can be used to produce antibodies specific to PbcA. The term "specific" is used to mean that when the antibodies to PbcA or fragments thereof are used in Western blots containing PbcA or a PbcA fragment or peptide under standard Western Blot assay conditions, the antibodies recognize only PbcA or its degradation or truncated products. Purified PbcA or fragments, preferably of at least about 15 amino acids in length, can be used to inject laboratory animals for the production of polyclonal and monoclonal antibodies to PbcA. Those skilled in the art will recognize that the methods for producing polyclonal antibodies and for producing monoclonal

antibodies are known in the art and include the methods disclosed by Harlow et al. (cited below). Purified antibodies can be generated using the isolated protein of this invention, or fragments thereof, without undue experimentation. The antibodies are useful in *in vitro* assays to test for the presence or absence of PbcA protein and to test for the ability of the antibodies to block C3 binding as well as in *in vivo* assays to test for the ability of the antibodies to provide passive protection against pneumococcal infection, whether local or systemic. Antibody fragments and chimeric antibodies can be used and the antibodies include at least one variable domain from an antibody specifically recognizing PbcA.

Moreover, purified proteins, peptides and polypeptides from PbcA can be injected into animals, and later humans, to produce an antibody response to *S. pneumoniae*. As used herein, the terms proteins, peptides and polypeptides are used interchangeably. Therefore, for purposes of this application and as used in the claims, a protein refers to proteins, protein fragments, peptides and polypeptides. Methods for introducing the protein, peptide or polypeptide fragment of PbcA to a mammal with an appropriate adjuvant, if necessary, are known in the art. Moreover, all or a portion of PbcA can be produced as an isolated protein or as a recombinant protein. Recombinant proteins can include all or a part of PcbA or can be formed as a chimeric protein. As used herein, the term "chimeric protein" refers to a recombinant protein including all or at least 15 amino acids of PcbA and amino acid sequence from at least one other protein positioned amino to, carboxy to, or on either side of the PbcA-derived amino acid. The 15 amino acids of PbcA of the chimeric protein are preferably unique to *S. pneumoniae*-derived proteins. Since patients with cleared *S. pneumoniae* infection have convalescent antibodies recognizing PbcA, it is known that the immune system can mount an immune response that includes the production of antibodies to PbcA. Ultimately, PbcA can be used as an immunogen for a pneumococcal vaccine. Further, based on the findings of these studies, antibody produced from the isolated protein, PbcA, can be used in Western Blot analyses to determine whether or not a particular *S. pneumoniae*

strain's virulence or avirulence has been associated with the presence or absence of PbcA.

As demonstrated in Example 4, PbcA is implicated in pneumococcal virulence (see Example 4). In this example an avirulent 23F pneumococcal isolate was inoculated into the ears of chinchillas. Virulent pneumococcal isolates typically cause otitis media and the influx of leukocytes after injection at concentrations as low as 1×10^2 colony forming units (cfu). In contrast, the avirulent 23F strain was inoculated at concentrations of up to about 1×10^7 cells and even at that level was unable to cause otitis media or inflammation (Giebink et al. *J. Infect. Dis.* 1993; 167:347-355). As discussed in Example 4, studies disclosed here indicate that the avirulent strain reported by Giebink lacked detectable PbcA protein by Western blot. To further study the implications of PbcA on virulence, an insertion/duplication mutation of PbcA was prepared for further study (see Example 8).

The present invention provides a detailed method of purification of PbcA and studies indicat that PbcA is: (a) immunogenic in man; (b) conserved within the mass range of about 90 kDa to about 110 kDa (as observed on a 15% SDS-PAGE gel) among a variety of pneumococcal serotypes (that is, PbcA has been identified in a variety of serotypes); and (c) absent in an avirulent 23F strain that is incapable of causing otitis media in a chinchilla model.

To demonstrate that PbcA can be detected in an ELISA assay, PbcA coated wells of an ELISA plate were incubated with antibody prepared to purified PbcA from *S. pneumoniae* strain CP1200.

To demonstrate the ability of PbcA to bind human C3, purified PbcA from CP1200 was used to coat ELISA plates and the coated protein was incubated with methylamine-treated human C3. The binding of methylamine-treated human C3 was assessed using antibody to human C3 conjugated with horseradish peroxidase (See Example 7). Figure 2 is a graph illustrating the ability of PbcA to bind increasing concentrations of human C3. Specific antibody to PbcA blocked the binding of purified human C3 to PbcA in a dose-dependent fashion. These experiments are detailed in Example 7 and in Figure

3. The Western blot (Figure 3a) demonstrates that the IgG fraction of antibody to PbcA blocked the binding of human C3 to PbcA immobilized on nitrocellulose. The graph (Figure 3b) demonstrates that affinity-purified antibody to PbcA blocked the binding of human C3 to PbcA immobilized on ELISA plates.

A PbcA insertion construct was prepared to interrupt a *pbc*A gene in CP1200 and R6x. An exemplary protocol is provided in Example 8 for producing an exemplary construct suitable for homologous recombination. Figure 4 illustrates a preferred method for preparing an insertion construct to inactivate the *pbc*A gene in *S. pneumoniae*. In these methods, a gene encoding PbcA (a non-native C3 binding protein) is introduced into a cell containing PbcA. The non-native C3 binding protein is homologous to the PbcA in the cell to facilitate homologous recombination and the production of insertional mutations. The presence of the insertion mutant was confirmed by Southern blot for a number of mutated R6X and CP1200 strains following transformation of these strains with the mutating construct pBV004 (see Figure 4). Mutated strains were tested by Southern blot using a 1.5 kb fragment of *pbc*A or vector pVA891 probes. Wild type strain BD23 was not transformed in these studies. Results of the Southern blot experiments are provided in Figure 5.

The insertion mutants were tested for their ability to produce PbcA protein using either human C3 or antibody to PbcA (Figure 4). Results indicated that no protein approximating the size of PbcA was detectable in supernatants from the *S. pneumoniae* strains tested on Western blot. Further, neither C3 nor antibody to PbcA bound on the Western blots. Therefore, the mutants are useful to assess functional aspects of PbcA and to serve as a negative control for a variety of PbcA-related experiments.

As noted above, PbcA binds noncovalently to C3. *S. pneumoniae* is generally cleared from the body by the covalent opsonic deposition of C3b on the pneumococcal capsule or cell wall, followed by phagocytosis via C3 receptors or Fc receptors on neutrophils or monocytes/macrophages. Nonopsonic (i.e. non-covalent) binding of C3 by pneumococcal surface proteins,

such as PbcA, suggests a mechanism whereby pneumococci can evade opsonization. Without intending to limit the scope of this invention, it is possible that PbcA protein from *S. pneumoniae* could bind C3 *in vivo* and reduce the amount of C3 available for opsonization. Antibodies to PbcA could block the C3-binding effect and restore the opsonic activity of C3 in plasma.

PbcA can be used in *in vitro* assays to assess the effect of PbcA on cells. For example, in Example 6, purified PbcA was added to cell cultures to study the role of PbcA in *S. pneumoniae* pathogenesis. PbcA can be tested for its toxicity on a variety of cells as well as tested in *in vivo* models for toxicity. The results of Example 6 indicated that PbcA, as an intact protein, was toxic to pulmonary epithelial cells and that it stimulated production of the cytokine IL-8 from the epithelial cells. IL-8 is a cytokine that, among other things, stimulates neutrophil migration. Increased concentrations of neutrophils are observed in the lungs of patients with significant *S. pneumoniae* infection in the lung passageways. Neutrophils and other white blood cells produce a variety of degradatory enzymes that damage lung tissue in *S. pneumoniae* infection and, based on these studies, lung damage can be the result of enzymatic release from white blood cells, PbcA, or both.

S. pneumoniae can colonize the nasopharynx, infect the lung and ultimately disseminate to the blood, meninges or other sites. Antibody to PbcA can be tested for its ability to reduce the toxicity of the organism to the lung tissue and antibody to PbcA can be tested for its ability to prevent PbcA binding to C3 and to permit C3 to remain available during *S. pneumoniae* infection. Similarly, peptides and polypeptides to PbcA can be administered to mammals and used in studies to assess the ability of the immune system to produce antibody to limit *S. pneumoniae* infection.

Animal models to study *S. pneumoniae* pathogenesis are known in the art. These include, but are not limited to, the chinchilla model for otitis media, the infant rat model for *S. pneumoniae* colonization and bacteremia, the mouse model for colonization, bacteremia and meningeal infection and a rabbit model for studying infection in the meninges. Those of ordinary skill in the art

will recognize that antibodies to PbcA (whether exogenously administered or the product of immunization with all or a part of PbcA) can be tested in these models for their ability to limit or inhibit *S. pneumoniae* infection. Exogenously (i.e., passively) administered antibody can be given through a variety of parenteral routes including, but not limited to, intravenous administration or administration to the air passages of a mammal such as a mouse, chinchilla, rat, rabbit or human.

All references and publications cited herein are expressly incorporated by reference into this disclosure. Particular embodiments of this invention will be discussed in detail and reference has been made to possible variations within the scope of this invention. There are a variety of alternative techniques and procedures available to those of skill in the art which would similarly permit one to successfully perform the intended invention.

Example 1
Identification of PbcA from *S. pneumoniae*

10 ml of *S. pneumoniae* strain CP1200 (obtained from D.A. Morrison, University of Illinois, Champagne-Urbana, Illinois and described in Havarstein LF, et al. *Proc. Natl. Acad. Sci. (USA)* 1995;92:11140-11144) was grown to exponential phase ($O.D._{620} \cong 0.3$) in Todd Hewitt broth (Fisher, Pittsburgh, PA) or in a synthetic medium ($O.D._{620} \cong 0.15$, media described by Sicard, A.M. *Genetics* 1964 59:31-44). Pneumococcal cells were pelleted and the supernatant was removed and precipitated with a final concentration of 10% trichloroacetic acid (TCA) at 4°C overnight and samples were electrophoresed on 15% sodium dodecyl sulfate polyacrylamide gels (SDS-PAGE) under non-reducing conditions. Pneumococcal proteins separated by electrophoresis were then transferred to nitrocellulose for Western blotting. After blocking of the Western blot according to standard protocols (Harlow , et al. *Antibodies; A Laboratory manual*. Cold Spring Harbor, NY; Cold Spring harbor Laboratory Press, 1988; 471-510) the blot was incubated with 10 mls of binding buffer containing about 2 µg/ml of purified human C3 (Hostetter et al. *J. Infect. Dis.*

1994; 150:653-661), labeled with biotin. The blot was washed and incubated with a 1:20,000 dilution of HRP-avidin for 60 minutes at room temperature and developed using the Supersignal™ system (Pierce, Rockford, IL) according to manufacturer's instructions. Purified human C3 bound to a band of about 90 kDa under non-reducing and reducing conditions on a 15% SDS-PAGE gel. Similarly, a protein of about 90 kDa was detected when cells were lysed in 5% SDS at room temperature for 30 minutes and following centrifugation, the supernatant was separated on a 15% SDS-PAGE gel.

Experiments were repeated with the substitution of 2 μg/ml biotinylated C3 after treatment with methylamine (using the methods disclosed in Hostetter, et al. *J. Infect. Dis.* 1984; 150:653-661) to demonstrate that the 90 kDa protein in pneumococcal supernatants and lysates could bind non-opsonic forms of C3. Results again indicated that a 90 kDa protein was recognized by C3. A variety of pneumococcal strains were also tested. C3 bound to PbcA identified in Western blots using supernatants from a variety of virulent pneumococcal strains (serotypes 1,3 [4 strains], 4, 14, 19F)

Example 2
Purification of PbcA

S. pneumoniae CP1200 was grown to mid-exponential phase in 4 liters of Todd Hewitt broth at 37°C. Pneumococcal cells were pelleted by centrifugation at 10,000 x g for 10 minutes. Proteins in the supernatant were precipitated with a final concentration of 10% trichloroacetic acid at 4 °C overnight. The precipitate was resuspended in 40 mls of 100 mM Tris and the final pH adjusted to 7.0. The resuspended proteins were chromatographed on a 1.2 ml column of Thiopropyl Separaose 6B coupled by a disulfide bond to 4 mg methylamine treated human C3. The column was then washed with 40 ml of 100 mM Tris-HCl, pH 7.0 containing 0.5 M NaCl. PbcA was eluted from the column with 20% ethanol in the Tris-HCl/NaCl wash buffer. 1 ml. fractions were collected and analyzed by SDS-PAGE and C3 binding assay. A protein of about 90 kDa

to about 110 kDa (+/- 5 kDa) eluted from the C3 affinity column in fractions 2-10.

Multiple eluates from sequential purifications of PbcA from strain 1200 were pooled by precipitation with 90% ethanol to obtain sufficient sample for sequencing and further protein studies. Approximately 80 picomoles of the protein were subjected to amino terminal analysis and tryptic digestion for internal peptide sequencing at the Harvard Microchemical Facility (Cambridge, MA).

Example 3
Immunogenicity of PbcA

Immunogenicity of PbcA was assessed by growing 10 ml of *S. pneumoniae* strain CP1200 to exponential phase in Todd Hewitt broth, pelleting the cells, and precipitating proteins from the supernatant with 10% TCA overnight at 4°C. The next day, supernatant proteins were electrophoresed on 15% SDS-PAGE, transferred to nitrocellulose, blocked with skim milk in a standard protocol (Harlow, et al. *supra*) and incubated with a 1:10,000 dilution of acute or convalescent serum from a patient infected with *S. pneumoniae* (Dr. E. Janoff, Minneapolis VA Hospital, Minneapolis, MN). The blot was washed according to methods disclosed in Harlow et al. and incubated with a 1:50,000 dilution of commercially purchased goat anti-human IgG conjugated to horseradish peroxidase (Chemicon, Temicula, CA). The blot was washed and developed with the Supersignal™ system according to manufacturer's instructions.

Western blots from these studies demonstrated that acute serum did not contain IgG antibodies to any pneumococcal proteins released into the supernatant but that convalescent serum contained IgG antibodies that recognized a protein of 90 kDa, consistent with the mass of PbcA. These experiments confirmed that PbcA elicited an immune response in humans recovering from *S. pneumoniae* infection and indicated that PbcA is recognized by the human immune system.

Example 4
PbcA is implicated in *S. pneumoniae* virulence

Virulent pneumococcal isolates typically cause otitis media and influx of leukocytes after inoculation in concentrations as low as 1 x 10^2 (Giebink et al. *J. Infect. Dis.* 1993; 167:347-355). Giebink et al. reported that an avirulent serotype, type 23F, was inoculated into the ears of chinchillas and that inoculum at concentrations of less than about 1 x 10^7 cells was unable to cause otitis media or inflammation.

Both the type 23F avirulent strain (GD 23, *supra*) and a type 23F virulent strain (BD23, Dr. Steve Pelton, Boston City Hospital, Boston, MA) were grown to mid-exponential phase in Todd Hewitt broth, the cells were pelleted, and supernatant proteins were precipitated in a final TCA concentration of 10% overnight at 4°C. The following day, the precipitate was resuspended in about 1 ml Tris, neutralized to pH 7.0, electrophoresed on 15% SDS-PAGE and then transferred to nitrocellulose. Incubation of the nitrocellulose membrane with 2 μg/ml of biotinylated, methylamine-treated C3 in binding buffer and development of the Western blot with avidin conjugated to horseradish peroxidase detected a PbcA band from a 15% SDS-PAGE gel in both cell lysates and supernatants from the virulent type 23F, but PbcA was completely absent in cell lysates and supernatants from the avirulent type 23F. In place of PbcA, a smaller band of 33 kDa was identified. This band may represent a degradation product or a truncated version of PbcA.

Example 5
Isolation of nucleic acid encoding PbcA

Degenerate oligonucleotides were obtained from a commercial supplier based on the sequence of SEQ ID NO:1 and SEQ ID NO:2. The oligonucleotides were used to amplify a 1500 bp sequence from CP1200 genomic DNA as template using standard polymerase chain reaction technology. Template, primers, and buffer were added for one 5-minute cycle at 94°C. Then dNTP's and Taq polymerase were added for 30 cycles, as follows: Double

stranded DNA was denatured for 1 min at 94°C, annealed for 1 min at 50°C, and extended for 2 min at 72°C. Final extension was completed in one 8-minute cycle at 72°C. The 1500 bp sequence was random primer labeled using a commercial kit and the sequence was used to screen the CP1200 genomic DNA library prepared under our direction by Stratagene (LaJolla, CA). Hybridization was performed under at least moderate stringency conditions and a variety of hybridization methods are provided in Sambrook et al. (*Molecular Cloning, A Laboratory Manual*, 1989 Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY). Exemplary hybridization conditions used were 65°C hybridization in 6X SSC (1M NaCl) using 5X Denhardt's, 20 mM sodium phosphate, 0.5% SDS and 100 μg/ml denatured sonicated salmon sperm DNA.

Three clones were identified that hybridized to the 1500 bp genomic DNA fragment; two overlapping clones of 4.3 and 5.3 kb, respectively and a third clone of 2.5 kb. Restriction mapping and hybridization studies suggested that the 5.3 kb clone contained an open reading frame encompassing the oligonucleotides from SEQ ID NO:1 and SEQ ID NO:2. A 3.2 kb fragment remaining after HindIII digestion and religation of the 5.3 kb clone is sequenced.

Example 6
Stimulation of IL-8 Production from Pulmonary Epithelial Cells in Response to PbcA

Monolayers of type II pulmonary epithelial cells (A549, American Type Culture Collection (ATCC), Rockville, MD) were incubated for 4 hours in culture supernatant (at least about 15μl supernatant diluted in 1 ml. serum free media (50% Hams F12/50% PBS) from exponentially growing *S. pneumoniae* (Pn) laboratory strain (CP1200, *supra)* or clinical isolates grown in Todd Hewitt broth (*supra*). Incubation of the epithelial cell monolayer in the culture supernatant induced the release of about 1169 pg/ml IL-8 while incubation in media alone without culture supernatant and incubation in pneumococcal growth medium alone without culture supernatant was significantly less ($p < 0.0001$, see table below).

Supernatants from a virulent *S. pneumoniae* clinical isolate, 23F, were also effective at stimulating IL-8 production from pulmonary epithelial cells (about 1,495 pg/ml IL-8). Supernatants from a 23F strain that did not produce otitis media in an animal model did not induce IL-8 release *in vitro*.

Increasing the time for which CP 1200 supernatants were incubated with the epithelial monolayer from 4 to 24 hours resulted in an increase in IL-8 production to approximately two times the level of IL-8 obtained after a 4 hour incubation.

SDS-PAGE analysis of proteins in CP1200 *S. pneumoniae* supernatants indicated that there a variety of proteins in the supernatants including five discrete bands correlating to identifiable proteins from *S. pneumoniae* that had sizes of about 180 kDa, 90 kDa, 57 kDa 42 kDa and 24 kDa. All five proteins were present in supernatants from the virulent 23F strain during exponential growth, but a band of about 90 kDa to about 110 kDa was absent in supernatants from the avirulent 23F strain.

In identical assays, about 150-200 ng of affinity purified PbcA protein elicited at least about 1200 pg/ml IL-8 from pulmonary epithelial cells. This level was similar to that elicited using supernatants from CP1200 and virulent 23F strains after a 4-hour incubation. Thus, not only is PbcA potent in eliciting IL-8 from pulmonary epithelial cells.

The combined results from 11 separate assays for IL-8 production are provided below:

Stimulus	**IL-8 Release (pg/ml) mean ±S.E.**
Medium alone	580 ± 40
Pneumo supt. Avir 23F(GD23)	471 ±72
Pneumo supt. Vir23F(BD23)	1,138 ± 83
Pneumo supt. CP1200	1,169 ±121
200 ng PbcA(CP1200)	1,679 ± 113
150 ng PbcA(CP1200)*	1,201
*assayed once	

Rabbit polyclonal antibody prepared against PbcA was used to determine whether the toxicity and IL-8 production stimulated by PbcA was

inhibited by PbcA-specific antibody as confirmation that the effects observed in these studies were attributable to PbcA.

Example 7
PbcA Binds Human C3

C3/PbcA ELISA Binding Assay

Purified PbcA was bound to ELISA 8-well strips (Costar, Cambridge, MA) overnight in binding buffer at room temp at varying concentrations (.5 μg/well and 0 μg/well). Wells were blocked with 5% Milk, 0.05% Tween 20 in PBS and 0.02% azide for 2 hours, washed 3 times with 0.05% Tween 20 in PBS and C3 was added (5 μg/well) in antibody diluent buffer (1% BSA, 1% Tween 20 in PBS) for 2 hours at 37°C. Plates were washed 3 times and incubated with HRP-conjugated goat anti-human C3 (1μg/ml) for 1 hour at 37°C. Plates were washed 3 times, developed with OPD (Zymed protocol) for 30 minutes, and the absorbence was read at A_{490} on an ELISA plate reader.

A standard curve was prepared to assess saturation of PbcA on the plate. Serial dilutions of PbcA protein were added (from 500 ng/well to 0 ng/well) to microtiter wells. Plates were incubated for 1 hour at 37°C with anti-PbcA (1:100), washed, incubated with HRP-conjugated goat anti-rabbit (1:2000) for 1 hour at 37°C. Plates were washed and developed with a 1:10 dilution of OPD developing buffer for 20 minutes at room temperature as per manufacturer's instructions and read at A_{490} on an ELISA plate reader.

Binding of C3 was determined by the immobilization of 500 ng PbcA/well. C3 was added in serial dilutions from 10 μg/well to 0 μg/well and incubated with immobilized PbcA for 2 hours at 37°C. Goat HRP-conjugated anti-human C3 (1 μg/well) was added and the absorbance was assessed. Figure 2 shows a dose-response curve for the binding of increasing amounts of C3 to constant amounts of immobilized PbcA.

Results indicated (see Figure 2) that increasing concentrations of C3 bound to 0.5 μg/well PbcA on the plate and saturated with a plateau of 6μg/well.

Initial blocking experiments were done using the IgG fraction of anti-PbcA polyclonal antibodies (2 to 8-fold molar excess over 500 mg PbcA added per well). The IgG fraction of serum from unimmunized rabbits served as a control. After PbcA protein was bound to the plate overnight and blocked for 2 hours, anti-PbcA antibodies were added for 1 hour at 37°C prior to the addition of C3. Plates were washed, 6 μg C3 was added to each well for 2 hours at 37°C. Plates were then incubated with goat HRP-conjugated anti-human C3 for 1 hour at 37°C then washed and developed for 4 minutes.

Both the IgG fraction of anti-PbcA antibodies and affinity purified antibodies inhibited C3 binding up to 50% while control IgG antibodies did not inhibit C3 binding. Figure 3a (Western blot) demonstrates that the IgG fraction of anti-PbcA blocked the binding of human C3 to PbcA immobilized on nitrocellulose. Figure 3b demonstrates that affinity-purified anti-PbcA blocked the binding of C3 to PbcA by 40-50%, while those antibodies remaining in rabbit serum after the removal of anti-PbcA antibodies had no blocking effect.

Example 8
Production of *pbc*A Insertion/Duplication Mutants of *S. Pneumoniae*

Generation of a *pbc*A insertion/duplication construct

To generate an insertion/duplication mutant of *pbc*A, a 761 bp fragment was isolated from pDF145 (a plasmid containing *pbc*A) first by digestion with ApoI (New England Biolabs) followed by the addition of DNA Polymerase I, Large (Klenow) Fragment to create a blunt end, and second by digestion with NheI to create a 5' overhang sticky end compatible with XbaI in the vector pVA891. This vector is a streptococcal/*E. coli* shuttle vector that replicated in *E. coli* but not in *S. pneumoniae* (from Dr. Gary Dunny, Dept. Microbiology, University of Minnesota, Minneapolis, MN and described in Macrina FL, et al. *Gene* 25:145-150, 1983). The 761 bp fragment was isolated on

a 0.7% agarose gel. Concurrently, the vector pVA891 was digested with ClaI and treated with Klenow to create a blunt end, followed by digestion with XbaI to create a 5' sticky end. The 761 bp fragment of *pbc*A was ligated into the pneumococcal vector pVA891 and transformed into competent DH5α *E. coli*. Tranformants were selected by resistance to chloramphenicol and plasmid DNA from 12 clones was cut with EcoRI and SpeI (the 761 fragment introduced a SpeI site). This confirmed the presence of the insert DNA in clones 2-12 and the absence of the insert in the control vector clone. Clone 4 was selected as construct pBV004 and DNA was transformed into unencapsulated laboratory strains R6x and CP1200 and selected by erythromycin (Erm) resistance.

Transformation of Knockout Construct for Insertion/Duplication Mutagenesis

Pneumococcal strains CP1200 and R6x were grown to OD_{550}= 0.2 and stored as frozen stock aliquots. Pneumococcal cultures were diluted 1:100 to 0.002 and grown to OD_{550}= 0.02. Competence was induced by the Morrison CSP protocol (Haverstein, L. et al. *Proc. Natl. Acad. Sci.* (USA) 92:11140-11144, 1995). To 100µl of cells, competence stimulating peptide (CSP) 100ng, was added along with 500ng to about 1 µg of construct pBV004 DNA. Cells were incubated for 30-40 minutes in a 37°C water bath with aeration to maintain constant temperature. After transformation, cells were diluted 1:10 in one ml total volume of THB+Y. DNase I (10µg/ml) was added and cells were incubated an additional 90 minutes to allow integration. The transformation mixture was diluted 1:10 and 100µl of cells were plated in a 4-layer agar overlay procedure as follows: first overlay, 3 mls THB agar; second overlay, 1.5 mls THBY + 1.5 mls THB agar + 100 µl transformation mixture, incubate 1 hour at 37° C; third overlay, 3 mls THB agar; fourth overlay, 3 mls THB agar + .05µg/ml Erythromycin.

Genomic DNA was isolated from wild type R6x, CP1200 and BD23(virulent 23F strain *supra*) strains as well as from insertional mutants in R6x and CP1200. Genomic DNA was digested with EcoRI (which does not

cut within the *pbc*A gene) and electrophoresed in 0.7% agarose, blotted onto nylon membrane and Southern blot performed with digoxygenin-labeled probes.

Referring to Figure 5, on the left are the results of a the Southern blot using the 1.5 kb fragment of *pbc*A as the probe. The probe hybridized with wild-type pneumococcal DNA in a single band and hybridized with two bands in the mutants R62r, R63r, R6x5 and CP1r, as expected.

On the right, the same blot was stripped and reprobed with pVA891 vector DNA. Wild-type DNA does not hybridize with pVA891, while the vector probe hybridizes at the same two bands indicating that the insertion mutagenesis was within the *pbc*A gene.

To confirm that the insertion mutated the gene and disrupted protein production, a Western blot analysis was performed (see Figure 6).

On the left, trichloroacetic acid (TCA) precipitated supernatants from strain R6x and the R62r mutant were electrophoresed on a 7.5% SDS-PAGE gel and blotted with either anti-PbcA antibodies and HRP-goat- anti-rabbit antibodies for detection of PbcA or with biotinylated C3 and HRP-avidin for detection of C3 binding. Insertional mutants did not synthesize PbcA and failed to bind to anti-PbcA antibodies or biotinylated C3. Absence of protein and lack of C3 binding was seen in the mutant compared to wild-type. The PbcA protein from R6x is larger than that of CP1200.

The original CP1200 detected a 90 kDa protein and had a truncated choline binding region due to premature termination. The R6x PbcA protein was about 105 kDa and appears to contain at least 9 choline binding repeats.

It will be appreciated by those skilled in the art that while the invention has been described above in connection with particular embodiments and examples, the invention is not necessarily so limited and that numerous other embodiments, examples, uses, modifications and departures from the embodiments, examples and uses may be made without departing from the inventive scope of this application.

SEQUENCE LISTING

(1) GENERAL INFORMATION:

(i) APPLICANT: REGENTS OF THE UNIVERSITY OF MINNESOTA, ET AL.

(ii) TITLE OF INVENTION: C3 BINDING PROTEIN OF STREPTOCOCCUS PNEUMONIAE

(iii) NUMBER OF SEQUENCES: 12

(iv) CORRESPONDENCE ADDRESS:
(A) ADDRESSEE: MUETING, RAASCH & GEBHARDT, P.A.
(B) STREET: 119 North Fourth Street, Suite 203
(C) CITY: Minneapolis
(D) STATE: Minnesota
(E) COUNTRY: USA
(F) ZIP: 55401

(v) COMPUTER READABLE FORM:
(A) MEDIUM TYPE: Floppy disk
(B) COMPUTER: IBM PC compatible
(C) OPERATING SYSTEM: PC-DOS/MS-DOS
(D) SOFTWARE: PatentIn Release #1.0, Version #1.30

(vi) CURRENT APPLICATION DATA:
(A) APPLICATION NUMBER: PCT/US97/20586
(B) FILING DATE: 12-NOV-1997
(C) CLASSIFICATION:

(vii) PRIORITY APPLICATION DATA:
(A) APPLICATION NUMBER: 60/029,444
(B) FILING DATE: 12-NOV-1996
(C) CLASSIFICATION:

(A) APPLICATION NUMBER: 60/038,086
(B) FILING DATE: 18-FEB-1997
(C) CLASSIFICATION:

(A) APPLICATION NUMBER: 60/059,368
(B) FILING DATE: 19-SEP-1997
(C) CLASSIFICATION:

(A) APPLICATION NUMBER: 60/062,473
(B) FILING DATE: 16-OCT-1997
(C) CLASSIFICATION

(viii) ATTORNEY/AGENT INFORMATION:
(A) NAME: MCCORMACK, MYRA H.
(B) REGISTRATION NUMBER: 36,602
(C) REFERENCE/DOCKET NUMBER: 110.00430201

(ix) TELECOMMUNICATION INFORMATION:
(A) TELEPHONE: 612-305-1225
(B) TELEFAX: 612-305-1228

(2) INFORMATION FOR SEQ ID NO:1:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

Thr Glu Asn Glu Gly Ser Thr Gln Ala Ala Thr Ser Ser Asn Met Ala
1 5 10 15

Lys Thr Glu His
20

(2) INFORMATION FOR SEQ ID NO:2:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 15 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

Glu Lys Pro Ala Glu Gln Pro Gln Pro Ala Pro Ala Thr Gln Pro
1 5 10 15

(2) INFORMATION FOR SEQ ID NO:3:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 17 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

Ser Ser Asp Ser Ser Val Gly Glu Glu Thr Leu Pro Ser Ser Ser Leu
1 5 10 15

Lys

(2) INFORMATION FOR SEQ ID NO:4:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 3023 base pairs

(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

GTAATACGAC TCACTATAGG GCGAATTGGG TACCGGGCCC CCCCTCGAGG TCGACGGTAT 60

CGATAAGCTT ATGCTTGTCA ATAATCACAA ATATGTAGAT CATATCTTGT TTAGGACAGT 120

AAAACATCCT AATTACTTTT TAAATATTCT TCCTGAGTTG ATTGGCTTGA CCTTGTTGAG 180

TCATGCTTAT GTGACTTTTG TTTTAGTTTT TCCAGTTTAT GCAGTTATTT TGTATCGACG 240

AATAGCTGAA GAGGAAAAGC TATTACATGA AGTTATAATC CCAAATGGAA GCATAAAGAG 300

ATAAATACAA AATTCGATTT ATATACAGTT CATATTGAAG TAATATAGTA AGGTTAAAGA 360

AAAAATATAG AAGGAAATAA ACATGTTTGC ATCAAAAAGC GAAAGAAAAG TACATTATTC 420

AATTCGTAAA TTTAGTATTG GAGTAGCTAG TGTAGCTGTT GCCAGTCTTG TTATGGGAAG 480

TGTGGTTCAT GCGACAGAGA ACGAGGGAAG TACCCAAGCA GCCACTTCTT CTAATATGGC 540

AAAGACAGAA CATAGGAAAG CTGCTAAACA AGTCGTCGAT GAATATATAG AAAAAATGTT 600

GAGGGAGATT CAACTAGATA GAAGAAAACA TACCCAAAAT GTCGCCTTAA ACATAAAGTT 660

GAGCGCAATT AAAACGAAGT ATTTGCGTGA ATTAAATGTT TTAGAAGAGA AGTCGAAAGA 720

TGAGTTGCCG TCAGAAATAA AAGCAAAGTT AGACGCAGCT TTTGAGAAGT TTAAAAAAGA 780

TACATTGAAA CCAGGAGAAA AGGTAGCAGA AGCTAAGAAG AAGGTTGAAG AAGCTAAGAA 840

AAAAGCCGAG GATCAAAAAG AAGAAGATCG TCGTAACTAC CCAACCAATA CTTACAAAAC 900

GCTTGAACTT GAAATTGCTG AGTTCGATGT GAAAGTTAAA GAAGCGGAGC TTGAACTAGT 960

AAAAGAGGAA GCTAAAGAAT CTCGAAACGA GGGCACAATT AAGCAAGCAA AAGAGAAAGT 1020

TGAGAGTAAA AAAGCTGAGG CTACAAGGTT AGAAAACATC AAGACAGATC GTAAAAAAGC 1080

AGAAGAAGAA GCTAAACGAA AAGCAGATGC TAAGTTGAAG GAAGCTAATG TAGCGACTTC 1140

AGATCAAGGT AAACCAAAGG GGCGGGCAAA ACGAGGAGTT CCTGGAGAGC TAGCAACACC 1200

TGATAAAAAA GAAAATGATG CGAAGTCTTC AGATTCTAGC GTAGGTGAAG AAACTCTTCC 1260

AAGCTCATCC CTGAAATCAG GAAAAAAGGT AGCAGAAGCT GAGAAGAAGG TTGAAGAAGC 1320

TGAGAAAAAA GCCAAGGATC AAAAAGAAGA AGATCGCCGT AACTACCCAA CCAATACTTA 1380

CAAAACGCTT GACCTTGAAA TTGCTGAGTC CGATGTGAAA GTTAAAGAAG CGGAGCTTGA 1440

ACTAGTAAAA GAGGAAGCTA AGGAACCTCG AGACGAGGAA AAAATTAAGC AAGCAAAAGC 1500

GAAAGTTGAG AGTAAAAAAG CTGAGGCTAC AAGGTTAGAA AACATCAAGA CAGATCGTAA 1560

AAAAGCAGAA GAAGAAGCTA AACGAAAAGC AGCAGAAGAA GATAAAGTTA AAGAAAAACC 1620

AGCTGAACAA CCACAACCAG CGCCGGCTAC TCAACCAGAA AAACCAGCTC CAAAACCAGA 1680

GAAGCCAGCT GAACAACCAA AAGCAGAAAA AACAGATGAT CAACAAGCTG AAGAAGACTA 1740

TGCTCGTAGA TCAGAAGAAG AATATAATCG CTTGACTCAA CAGCAACCGC CAAAAACTGA 1800

AAAACCAGCA CAACCATCTA CTCCAAAAAC AGGCTGGAAA CAAGAAAACG GTATGTGGTA 1860

CTTCTACAAT ACTGATGGTT CAATGGCAAC AGGATGGCTC CAAAACAACG GTTCATGGTA 1920

CTATCTAAAC GCTAATGGTG CTATGGCGAC AGGATGGCTC CAAAACAATG GTTCATGGTA 1980

CTATCTAAAC GCTAATGGTT CAATGGCAAC AGGATGGCTC CAAAACAATG GTTCATGGTA 2040

CTACCTAAAC GCTAATGGTG CTATGGCGAC AGGATAGCTC CAATACAATG GTTCATGGTA 2100

CTACCTAAAC AGCAATGGCG CTATGGCGAC AGGATGGCTC CAATACAATG GCTCATGGTA 2160

CTACCTCAAC GCTAATGGTG ATATGGCGAC AGGATGGCTC CAAAACAACG GTTCATGGTA 2220

CTACCTCAAC GCTAATGGTG ATATGGCGAC AGGATGGCTC CAATACAACG GTTCATGGTA 2280

TTACCTCAAC GCTAATGGTG ATATGGCGAC AGGTTGGGTG AAAGATGGAG ATACCTGGTA 2340

CTATCTTGAA GCATCAGGTG CTATGAAAGC AAGCCAATGG TTCAAAGTAT CAGATAAATG 2400

GTACTATGTC AATGGCTCAG GTGCCCTTGC AGTCAACACA ACTGTAGATG GCTATGGAGT 2460

CAATGCCAAT GGTGAATGGG TAAACTAAAC CTAATATAAC TAGTTAATAC TGACTTCCTG 2520

TAAGAACTTT TTAAAGTATT CCCTACAAAT ACCATATCCT TTCAGTAGAT AATATACCCT 2580

TGTAGGAAGT TTAGATTAAA AAATAACTCT GTAATCTCTA GCCGGATTTA TAGCGCTAGA 2640

GACTACGGAG TTTTTTTGAT GAGGAAAGAA TGGCGGCATT CAAGAGACTC TTTAAGAGAG 2700

TTACGGGTTT TAAACTATTA AGCCTTCTCC AATTGCAAGA GGGCTTCAAT CTCTGCTAGG 2760

GTGCTAGCTT GCGAAATGGC TCCACGGAGT TTGGCAGCGC CAGATGTTCC ACGGAGATAG 2820

TGAGGAGCGA GGCCGCGGAA TTCACGAACT GCGACGTTTT CTCCTTTGAG GTTAATCAAT 2880

CGTTTCAGGA ATTCCGGAAT TCCGGAATTC CGGAATTCCG GAATTCCGGA ATTCCTGCAG 2940

CCCGGGGGAT CCACTAGTTC TAGAGCGGCC GCCACCGCGG TGGAGCTCCA GCTTTTGTTC 3000

CCTTTAGTGA GGGTTAATTT CGA 3023

(2) INFORMATION FOR SEQ ID NO:5:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 1695 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: double
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

CGAAACGAGG GCACAATTAA GCAAGCAAAA GAGAAAGTTG AGAGTAAAAA AGCTGAGGCT 60

ACAAGGTTAG AAAACATCAA GACAGATCGT AAAAAAGCAG AAGAAGAAGC TAAACGAAAA 120

GCAGATGCTA AGTTGAAGGA AGCTAATGTA GCGACTTCAG ATCAAGGTAA ACCAAAGGGG 180

CGGGCAAAAC GAGGAGTTCC TGGAGAGCTA GCAACACCTG ATAAAAAAGA AAATGATGCG 240

AAGTCTTCAG ATTCTAGCGT AGGTGAAGAA ACTCTTCCAA GCTCATCCCT GAAATCAGGA 300

AAAAAGGTAG CAGAAGCTGA GAAGAAGGTT GAAGAAGCTG AGAAAAAAGC CAAGGATCAA 360

AAAGAAGAAG ATCGCCGTAA CTACCCAACC AATACTTACA AAACGCTTGA CCTTGAAATT 420

GCTGAGTCCG ATGTGAAAGT TAAAGAAGCG GAGCTTGAAC TAGTAAAAGA GGAAGCTAAG 480

GAACCTCGAG ACGAGGAAAA AATTAAGCAA GCAAAAGCGA AAGTTGAGAG TAAAAAAGCT 540

GAGGCTACAA GGTTAGAAAA CATCAAGACA GATCGTAAAA AAGCAGAAGA AGAAGCTAAA 600

ATGTTTGCAT CAAAAAGCGA AAGAAAAGTA CATTATTCAA TTCGTAAATT TAGTATTGGA 660

GTAGCTAGTG TAGCTGTTGC CAGTCTTGTT ATGGGAAGTG TGGTTCATGC GACAGAGAAC 720

GAGGGAAGTA CCCAAGCAGC CACTTCTTCT AATATGGCAA AGACAGAACA TAGGAAAGCT 780

GCTAAACAAG TCGTCGATGA ATATATAGAA AAAATGTTGA GGGAGATTCA ACTAGATAGA 840

AGAAAACATA CCCAAAATGT CGCCTTAAAC ATAAAGTTGA GCGCAATTAA AACGAAGTAT 900

TTGCGTGAAT TAAATGTTTT AGAAGAGAAG TCGAAAGATG AGTTGCCGTC AGAAATAAAA 960

GCAAAGTTAG ACGCAGCTTT TGAGAAGTTT AAAAAAGATA CATTGAAACC AGGAGAAAAG 1020

GTAGCAGAAG CTAAGAAGAA GGTTGAAGAA GCTAAGAAAA AAGCCGAGGA TCAAAAAGAA 1080

GAAGATCGTC GTAACTACCC AACCAATACT TACAAAACGC TTGAACTTGA AATTGCTGAG 1140

TTCGATGTGA AAGTTAAAGA AGCGGAGCTT GAACTAGTAA AAGAGGAAGC TAAAGAATCT 1200

CGAAAAGCAG CAGAAGAAGA TAAAGTTAAA GAAAAACCAG CTGAACAACC ACAACCAGCG 1260

CCGGCTACTC AACCAGAAAA ACCAGCTCCA AAACCAGAGA AGCCAGCTGA ACAACCAAAA 1320

GCAGAAAAAA CAGATGATCA ACAAGCTGAA GAAGACTATG CTCGTAGATC AGAAGAAGAA 1380

TATAATCGCT TGACTCAACA GCAACCGCCA AAAACTGAAA AACCAGCACA ACCATCTACT 1440

CCAAAAACAG GCTGGAAACA AGAAAACGGT ATGTGGTACT TCTACAATAC TGATGGTTCA 1500

ATGGCAACAG GATGGCTCCA AAACAACGGT TCATGGTACT ATCTAAACGC TAATGGTGCT 1560

ATGGCGACAG GATGGCTCCA AAACAATGGT TCATGGTACT ATCTAAACGC TAATGGTTCA 1620

ATGGCAACAG GATGGCTCCA AAACAATGGT TCATGGTACT ACCTAAACGC TAATGGTGCT 1680

ATGGCGACAG GATAG 1695

(2) INFORMATION FOR SEQ ID NO:6:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 564 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

Phe Asp Val Lys Val Lys Glu Ala Glu Leu Glu Leu Val Lys Glu Glu
1 5 10 15

Ala Lys Glu Ser Arg Asn Glu Gly Thr Ile Lys Gln Ala Lys Glu Lys
20 25 30

Val Glu Ser Lys Lys Ala Glu Ala Thr Arg Leu Glu Asn Ile Lys Thr
35 40 45

Asp Arg Lys Lys Ala Glu Glu Glu Ala Lys Arg Lys Ala Asp Ala Lys
50 55 60

Leu Lys Glu Ala Asn Val Ala Thr Ser Asp Gln Gly Lys Pro Lys Gly
65 70 75 80

Arg Ala Lys Arg Gly Val Pro Gly Glu Leu Ala Thr Pro Asp Lys Lys
85 90 95

Glu Asn Asp Ala Lys Ser Ser Asp Ser Ser Val Gly Glu Glu Thr Leu
100 105 110

Pro Ser Ser Ser Leu Lys Ser Gly Lys Lys Val Ala Glu Ala Glu Lys
115 120 125

Lys Val Glu Glu Ala Glu Lys Lys Ala Lys Asp Gln Lys Glu Glu Asp
130 135 140

Arg Arg Asn Tyr Pro Thr Asn Thr Tyr Lys Thr Leu Asp Leu Glu Ile
145 150 155 160

Ala Glu Ser Asp Val Lys Val Lys Glu Ala Glu Leu Glu Ile Val Lys
165 170 175

Glu Glu Ala Lys Glu Pro Arg Asp Glu Glu Lys Ile Lys Gln Ala Lys
180 185 190

Ala Lys Val Glu Ser Lys Lys Ala Glu Ala Thr Arg Leu Glu Asn Ile
195 200 205

Lys Thr Asp Arg Lys Lys Ala Glu Glu Glu Ala Lys Met Phe Ala Ser
210 215 220

Lys Ser Glu Arg Lys Val His Tyr Ser Ile Arg Lys Phe Ser Ile Gly
225 230 235 240

Val Ala Ser Val Ala Val Ala Ser Leu Val Met Gly Ser Val Val His
245 250 255

Ala Thr Glu Asn Glu Gly Ser Thr Gln Ala Ala Thr Ser Ser Asn Met
260 265 270

Ala Lys Thr Glu His Arg Lys Ala Ala Lys Gln Val Val Asp Glu Tyr
275 280 285

Ile Glu Lys Met Leu Arg Glu Ile Gln Leu Asp Arg Arg Lys His Thr
290 295 300

Gln Asn Val Ala Leu Asn Ile Lys Leu Ser Ala Ile Lys Thr Lys Tyr
305 310 315 320

Leu Arg Glu Leu Asn Val Leu Glu Glu Lys Ser Lys Asp Glu Leu Pro
325 330 335

Ser Glu Ile Lys Ala Lys Leu Asp Ala Ala Phe Glu Lys Phe Lys Lys
340 345 350

Asp Thr Leu Lys Pro Gly Glu Lys Val Ala Glu Ala Lys Lys Lys Val
355 360 365

Glu Glu Ala Lys Lys Lys Ala Glu Asp Gln Lys Glu Glu Asp Arg Arg
370 375 380

Asn Tyr Pro Thr Asn Thr Tyr Lys Thr Leu Glu Leu Glu Ile Ala Glu
385 390 395 400

Arg Lys Ala Ala Glu Glu Asp Lys Val Lys Glu Lys Pro Ala Glu Gln
405 410 415

Pro Gln Pro Ala Pro Ala Thr Gln Pro Glu Lys Pro Ala Pro Lys Pro
420 425 430

Glu Lys Pro Ala Glu Gln Pro Lys Ala Glu Lys Thr Asp Asp Gln Gln
435 440 445

Ala Glu Glu Asp Tyr Ala Arg Arg Ser Glu Glu Glu Tyr Asn Arg Leu
450 455 460

Thr Gln Gln Gln Pro Pro Lys Thr Glu Lys Pro Ala Gln Pro Ser Thr
465 470 475 480

Pro Lys Thr Gly Trp Lys Gln Glu Asn Gly Met Trp Tyr Phe Tyr Asn
485 490 495

Thr Asp Gly Ser Met Ala Thr Gly Trp Leu Gln Asn Asn Gly Ser Trp
500 505 510

Tyr Tyr Leu Asn Ala Asn Gly Ala Met Ala Thr Gly Trp Leu Gln Asn
515 520 525

Asn Gly Ser Trp Tyr Tyr Leu Asn Ala Asn Gly Ser Met Ala Thr Gly
530 535 540

Trp Leu Gln Asn Asn Gly Ser Trp Tyr Tyr Leu Asn Ala Asn Gly Ala
545 550 555 560

Met Ala Thr Gly

(2) INFORMATION FOR SEQ ID NO:7:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 18 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

GCACAACCAT CTACTCCA 18

(2) INFORMATION FOR SEQ ID NO:8:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 23 base pairs
(B) TYPE: nucleic acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: DNA (genomic)

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

GTACAGGAAT TCAGTATTAA CTA 23

(2) INFORMATION FOR SEQ ID NO:9:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 220 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

Asp Arg Trp Lys Gln Glu Asn Gly Met Trp Tyr Phe Tyr Asn Thr Asp
1 5 10 15

Gly Ser Met Ala Thr Gly Trp Leu Gln Asn Asn Gly Ser Trp Tyr Tyr
20 25 30

Leu Asn Ala Asn Gly Ala Met Ala Thr Gly Trp Leu Gln Asn Asn Gly
35 40 45

Ser Trp Tyr Tyr Leu Asn Ala Asn Gly Ser Met Ala Thr Gly Trp Leu
50 55 60

Gln Asn Asn Gly Ser Trp Tyr Tyr Leu Asn Ala Asn Gly Ala Met Ala
65 70 75 80

Thr Gly Leu Gln Tyr Asn Gly Ser Trp Tyr Tyr Leu Asn Ser Asn Gly
85 90 95

Ala Met Ala Thr Gly Trp Leu Gln Tyr Asn Gly Ser Trp Tyr Tyr Leu
100 105 110

Asn Ala Asn Gly Asp Met Ala Thr Gly Trp Leu Gln Asn Asn Gly Ser
115 120 125

Trp Tyr Tyr Leu Asn Ala Asn Gly Asp Met Ala Thr Gly Trp Leu Gln
130 135 140

Tyr Asn Gly Ser Trp Tyr Tyr Leu Asn Ala Asn Gly Asp Met Ala Thr
145 150 155 160

Gly Trp Val Lys Asp Gly Asp Thr Trp Tyr Tyr Leu Glu Ala Ser Gly
165 170 175

Ala Met Lys Ala Ser Gln Trp Phe Lys Val Ser Asp Lys Trp Tyr Tyr
180 185 190

Val Asn Gly Ser Gly Ala Leu Ala Val Asn Thr Thr Val Asp Gly Tyr
195 200 205

Gly Val Asn Ala Asn Gly Glu Trp Thr Lys His Tyr
210 215 220

(2) INFORMATION FOR SEQ ID NO:10:

(i) SEQUENCE CHARACTERISTICS:

(A) LENGTH: 217 amino acids

(B) TYPE: amino acid

(C) STRANDEDNESS: single

(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

Lys Gln Glu Asn Gly Met Trp Tyr Phe Tyr Asn Thr Asp Gly Ser Met
1 5 10 15

Ala Thr Gly Trp Leu Gln Asn Asn Gly Ser Trp Tyr Tyr Leu Asn Ala
20 25 30

Asn Gly Ala Met Ala Thr Gly Trp Leu Gln Asn Asn Gly Ser Trp Tyr
35 40 45

Tyr Leu Asn Ala Asn Gly Ser Met Ala Thr Gly Trp Leu Gln Asn Asn
50 55 60

Gly Ser Trp Tyr Tyr Leu Asn Ala Asn Gly Ala Met Ala Thr Gly Trp
65 70 75 80

Leu Gln Tyr Asn Gly Ser Trp Tyr Tyr Leu Asn Ser Asn Gly Ala Met
85 90 95

Ala Thr Gly Trp Leu Gln Tyr Asn Gly Ser Trp Tyr Tyr Leu Asn Ala
100 105 110

Asn Gly Asp Met Ala Thr Gly Trp Leu Gln Asn Asn Gly Ser Trp Tyr
115 120 125

Tyr Leu Asn Ala Asn Gly Asp Met Ala Thr Gly Trp Leu Gln Tyr Asn
130 135 140

Gly Ser Trp Tyr Tyr Leu Asn Ala Asn Gly Asp Met Ala Thr Gly Trp
145 150 155 160

Val Lys Asp Gly Asp Thr Trp Tyr Tyr Leu Glu Ala Ser Gly Ala Met
165 170 175

Lys Ala Ser Gln Trp Phe Lys Val Ser Asp Lys Trp Tyr Tyr Val Asn
180 185 190

Gly Ser Gly Ala Leu Ala Val Asn Thr Thr Val Asp Gly Tyr Gly Val
195 200 205

Asn Ala Asn Gly Glu Trp Thr Lys Pro
210 215

(2) INFORMATION FOR SEQ ID NO:11:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 220 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: protein

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

Leu Glu Thr Arg Asn Gly Met Trp Tyr Phe Tyr Asn Thr Asp Gly Ser
1 5 10 15

Met Ala Thr Gly Trp Leu Gln Asn Asn Gly Ser Trp Tyr Tyr Leu Asn
20 25 30

Ser Asn Gly Ala Met Ala Thr Gly Trp Leu Gln Asn Asn Gly Ser Trp
35 40 45

Tyr Tyr Leu Asn Ala Asn Gly Ser Met Ala Thr Gly Trp Leu Gln Asn
50 55 60

Asn Gly Ser Trp Tyr Tyr Leu Asn Ala Asn Gly Ala Met Ala Thr Gly
65 70 75 80

Trp Leu Gln Tyr Asn Gly Ser Trp Tyr Tyr Leu Asn Ser Asn Gly Ala
85 90 95

Met Ala Thr Gly Trp Leu Gln Tyr Asn Gly Ser Trp Tyr Tyr Leu Asn
100 105 110

Ala Asn Gly Asp Met Ala Thr Gly Trp Leu Gln Asn Asn Gly Ser Trp
115 120 125

Tyr Tyr Leu Asn Ala Asn Gly Asp Met Ala Thr Gly Trp Leu Gln Tyr
130 135 140

Asn Gly Ser Trp Tyr Tyr Leu Asn Ala Asn Gly Asp Met Ala Thr Gly
145 150 155 160

Trp Val Lys Asp Gly Asp Thr Trp Tyr Tyr Leu Glu Ala Ser Gly Ala
165 170 175

Met Lys Ala Ser Gln Trp Phe Lys Val Ser Asp Lys Trp Tyr Tyr Val
180 185 190

Asn Gly Ser Gly Ala Leu Ala Val Asn Thr Thr Val Asp Gly Tyr Gly
195 200 205

Val Asn Ala Asn Gly Glu Trp Thr Lys Pro Asn Ile
210 215 220

(2) INFORMATION FOR SEQ ID NO:12:

(i) SEQUENCE CHARACTERISTICS:
(A) LENGTH: 20 amino acids
(B) TYPE: amino acid
(C) STRANDEDNESS: single
(D) TOPOLOGY: linear

(ii) MOLECULE TYPE: peptide

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

Thr Gly Trp Lys Gln Glu Asn Gly Met Trp Tyr Phe Tyr Asn Thr Asp
1 5 10 15

Gly Ser Met Ala
20

WHAT IS CLAIMED IS:

1. An isolated protein comprising SEQ ID NO:1 and at least two choline binding repeats.

2. The protein of claim 1 wherein the protein is isolated from *S. pneumoniae*.

3. The protein of claim 1 wherein the protein binds human complement protein C3.

4. The protein of claim 1, wherein the protein is a recombinant protein.

5. The protein of claim 1 having a molecular weight as determined on a 15% polyacrylamide gel of between about 90 kDa to about 110 kDa.

6. The protein of claim 1 further comprising a proline rich region.

7. The protein of claim 1 further comprising SEQ ID NO:2 or SEQ ID NO:3.

8. An isolated protein comprising SEQ ID NO:6.

9. An isolated protein capable of binding to, but not cleaving or degrading, human complement C3 and wherein the protein comprises SEQ ID NO:1.

10. The protein of claim 9 wherein the protein is isolated from *S. pneumoniae*.

11. The protein of claim 9 further comprising a proline rich region.

12. The protein of claim 9 further comprising SEQ ID NO:2.

13. The protein of claim 9 having at least about 95% homology to a C3 binding protein from *S. pneumoniae*.

14. The protein of claim 9 having a molecular weight as determined on a 15% polyacrylamide gel of between about 90 kDa to about 110 kDa.

15. The protein of claim 9, wherein the protein is a recombinant protein.

16. A recombinant protein comprising SEQ ID NO:1, wherein the protein has a molecular weight as determined on a 15% polyacrylamide gel of between about 90 kDa to about 110 kDa.

17. The protein of claim 16 wherein the protein binds human complement protein C3.

18. The protein of claim 16 further comprising a proline rich region.

19. The protein of claim 17, wherein the protein does not cleave or degrade human complement protein C3.

20. A protein comprising amino acids 1-410 of SEQ ID NO:6.

21. A protein that binds, but does not cleave or degrade, human complement protein C3, wherein nucleic acid encoding the protein hybridizes to SEQ ID NO:4 under hybridization conditions of 6XSSC, 5X Denhardt, 0.5% SDS, and 100 μg/ml fragmented and denatured salmon sperm DNA hybridized overnight at 65°C and washed in 2X SSC, 0.1% SDS one time at room temperature for about 10 minutes followed by one time at, 65°C for about 15 minutes followed

by at least one wash in 0.2XSSC, 0.1% SDS at room temperature for at least 3-5 minutes.

22. The protein of claim 21 further comprising SEQ ID NO:1.

23. The protein of claim 21 further comprising SEQ ID NO:2 or SEQ ID NO:3.

24. The protein of claim 21 wherein the protein is at least 15 amino acids in length.

25. The protein of claim 21, wherein the protein is a recombinant protein.

26. The protein of claim 21 further comprising a proline rich region.

27. The protein of claim 21, wherein the protein is a synthetic peptide.

28. The protein of claim 21, wherein the protein is a peptide of at least 15 amino acids from SEQ ID NO:6.

29. Antibody capable of specifically binding to the protein of claim 21.

30. The antibody of claim 29, wherein the antibody is a monoclonal antibody.

31. The antibody of claim 29, wherein the antibody is a polyclonal antibody.

32. The antibody of 30, wherein the antibody is rodent-derived.

33. Nucleic acid encoding a protein comprising at least two choline binding domains and SEQ ID NO: 1.

34. The nucleic acid of claim 33 further comprising a proline rich region.

35. The nucleic acid of claim 33 isolated from an *S. pneumoniae* genome.

36. The nucleic acid of claim 33 capable of hybridizing to SEQ ID NO:4.

37. The nucleic acid of claim 33 wherein the protein encoded by the nucleic acid binds to human complement protein C3.

38. The nucleic acid of claim 33 in a nucleic acid vector.

39. The nucleic acid of claim 33 wherein the vector is an expression vector.

40. The nucleic acid of claim 38 in a cell.

41. The nucleic acid of claim 40 wherein the cell is a bacterium.

42. A bacterium expressing a recombinant protein according to claim 21.

43. The nucleic acid of claim 39 in a cell.

44. Protein produced by the nucleic acid of claim 43.

45. Isolated nucleic acid encoding a protein comprising SEQ ID NO:1 and a proline rich region wherein the protein encoded by the nucleic acid binds but does not cleave or degrade human complement C3.

46. An isolated nucleic acid fragment encoding an about 90 kDa to about 110 kDa protein with C3 binding activity, wherein the nucleic acid fragment has at least 80% homology to at least 500 bp from nucleic acids 1-1500 of SEQ ID NO:5.

47. An isolated nucleic acid fragment comprising base pairs 1-1500 of SEQ ID NO:5.

48. A method for isolating a C3 binding protein from a bacterium comprising the steps of:

obtaining a protein sample from a bacterium;

applying the sample to a solid support comprising methylamine treated complement protein C3;

washing the solid support; and

removing a C3 binding protein from the solid support in a solution comprising alcohol;

wherein the C3 binding protein does not cleave or degrade C3.

49. The method of claim 48 wherein the bacterium is *S. pneumoniae*.

50. The method of claim 48 wherein the bacterium is *E. coli*.

51. The method of claim 48 wherein the solid support comprises an affinity column.

52. The method of claim 48 wherein the alcohol is ethanol.

53. The method of claim 52 wherein the solution comprising alcohol is a buffer comprising 20% ethanol.

54. C3 binding protein preparable by the method of claim 48.

55. A method for producing an immune response to *S. pneumoniae* comprising the steps of:

administering a therapeutically effective amount of at least a portion of a protein to a mammal, wherein the protein binds but

does not cleave or degrade human complement protein C3 and, wherein nucleic acid encoding the protein hybridizes to SEQ ID NO:4 under hybridization conditions of 6XSSC, 5X Denhardt, 0.5% SDS, and 100 μg/ml fragmented and denatured salmon sperm DNA, hybridized overnight at 65°C and washed in 2x SSC, 0.1% SDS one time at room temperature for about 10 minutes followed by one time at 65°C for about 15 minutes followed by at least one wash in 0.2xSSC, 0.1% SDS at room temperature for at least 3-5 minutes; and
detecting an immune response to the protein.

56. The method of claim 55 wherein the protein is at least 15 amino acids in length.

57. The method of claim 55 wherein the protein is a chimeric protein.

58. The method of claim 55 wherein the protein comprises SEQ ID NO:1.

59. The method of claim 55 wherein the protein has a molecular weight on a 15% polyacrylamide gel of between about 90 kDa to about 110 kDa.

60. A method for reducing *S. pneumoniae* binding to C3 comprising the steps of:
administering a therapeutically effective amount of at least a portion of an antibody to a mammal, wherein the antibody specifically recognizes a C3 binding protein from *S. pneumoniae*, and wherein the C3 binding protein is a protein that binds, but does not cleave or degrade human complement protein C3.

61. The method of claim 60 wherein the antibody comprises at least one variable domain from a monoclonal antibody.

62. The method of claim 60 wherein the antibody is administered to the air passages of the mammal.

63. The method of claim 60 wherein the antibody is administered intravenously to a mammal.

64. A non-naturally occurring *S. pneumoniae* bacterium that does not express a detectable human complement C3 binding protein, wherein the C3 binding protein comprises SEQ ID NO:1.

65. The bacterium of claim 64 wherein the bacterium is the product of an insertion into the gene encoding the C3 binding protein.

66. The bacterium of claim 64 wherein the bacterium is a product of a deletion in the gene encoding the C3 binding protein.

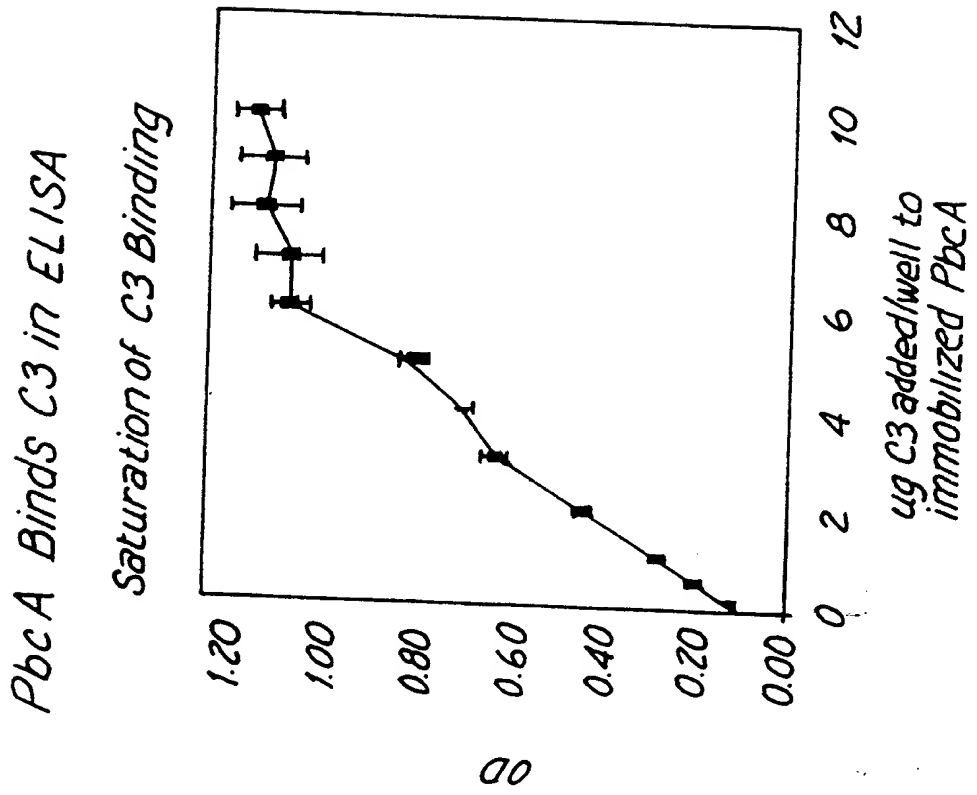
67. The bacterium of claim 64 produced by homologous recombination of the gene encoding the C3 binding protein gene with at least a portion of a non-native C3 binding protein gene.

68. The method of claim 67 wherein the gene encoding at least a portion of the non-native C3 binding protein comprises a mutation within the non-native gene.

```
                                                                  50
1
DRWKQENGMW YFYNTDGSMA TGWLQNNGSW YYLNANGAMA TGWLQNNGSW
---KQENGMW YFYNTDGSMA TGWLQNNGSW YYLNANGAMA TGWLQNNGSW
--LETRNGMW YFYNTDGSMA TGWLQNNGSW YYLNSNGAMA TGWLQNNGSW
1200
 r6x
 23f

                                                                 100
51
YYLNANGSMA TGWLQNNGSW YYLNANGAMA TG*LQYNGSW YYLNSNGAMA
YYLNANGSMA TGWLQNNGSW YYLNANGAMA TGWLQYNGSW YYLNSNGAMA
YYLNANGSMA TGWLQNNGSW YYLNANGSMA TGWLQYNGSW YYLNANGDMA
1200
 r6x
 23f

                                                                 150
101
TGWLQYNGSW YYLNANGDMA TGWLQNNGSW YYLNANGDMA TGWLQYNGSW
TGWLQYNGSW YYLNANGDMA TGWLQNNGSW YYLNANGDMA TGWLQYNGSW
TGWLQNNGSW YYLNANGDMA TGWLQNNGSW YYLNANGDMA TGWLQYNGSW
1200
 r6x
 23f

                                                                 200
151
YYLNANGDMA TGWVKDGDTW YYLEASGAMK ASQWFKVSDK WYYVNGSGAL
YYLNANGDMA TGWVKDGDTW YYLEASGAMK ASQWFKVSDK WYYVNGSGAL
YYLNANGDME TGWVKDGDTW YYLEASGAMK ASQWFKVSDK WYYVNGSGAL
1200
 r6x
 23F

                                                          223
201
AVNTTVDGYG VNANGEW*TK H*Y
AVNTTVDGYG VNANGEW*TK P--
AVNTTVDGYG VNANGEW*TK PNI
1200
 r6x
 23f
```

Fig. 1

Fig. 2


PbcA Binds C3 in ELISA
Saturation of C3 Binding
OD
1.20
1.00
0.80
0.60
0.40
0.20
0.00
0
2
4
6
8
10
12
ug C3 added/well to immobilized PbcA

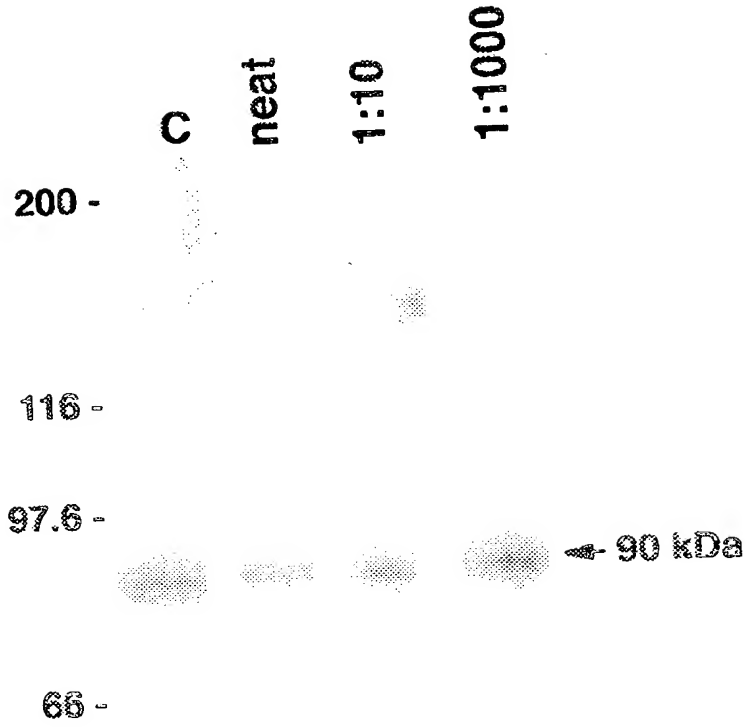


Fig. 3a

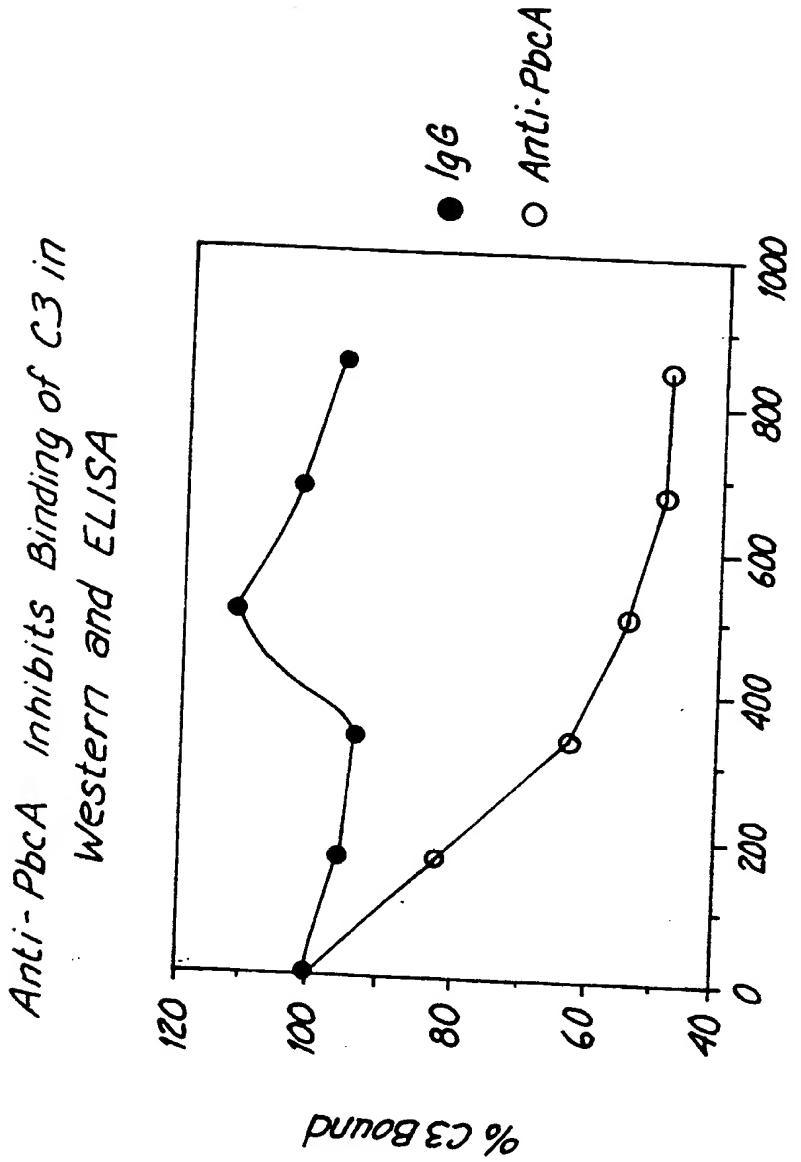

Anti-PbcA Inhibits Binding of C3 in Western and ELISA
● IgG
○ Anti-PbcA
% C3 Bound
120
100
80
60
40
0
200
400
600
800
1000
ug Antibody

Fig. 3b

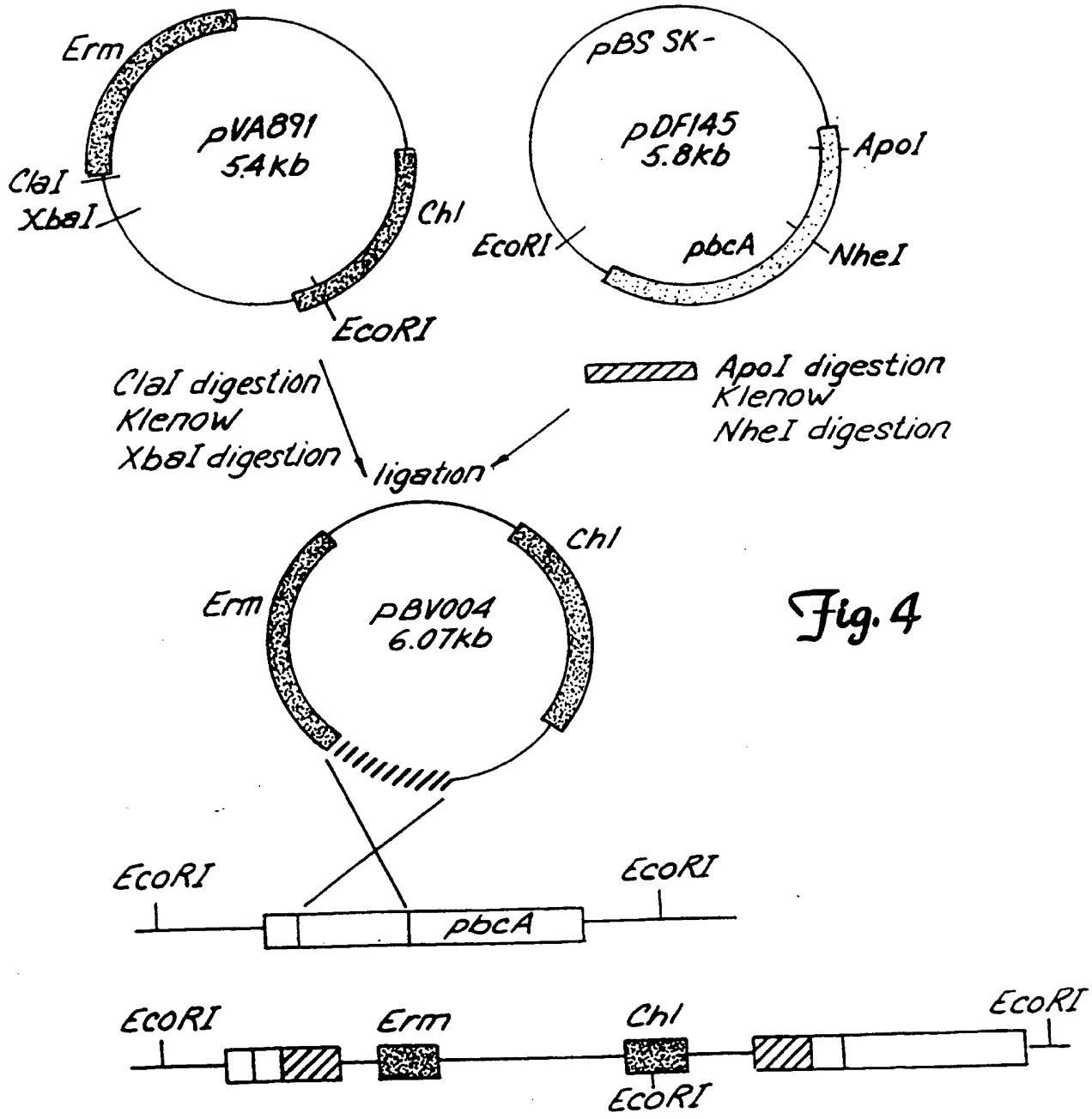


Fig. 4

Insertion-duplication mutagenesis

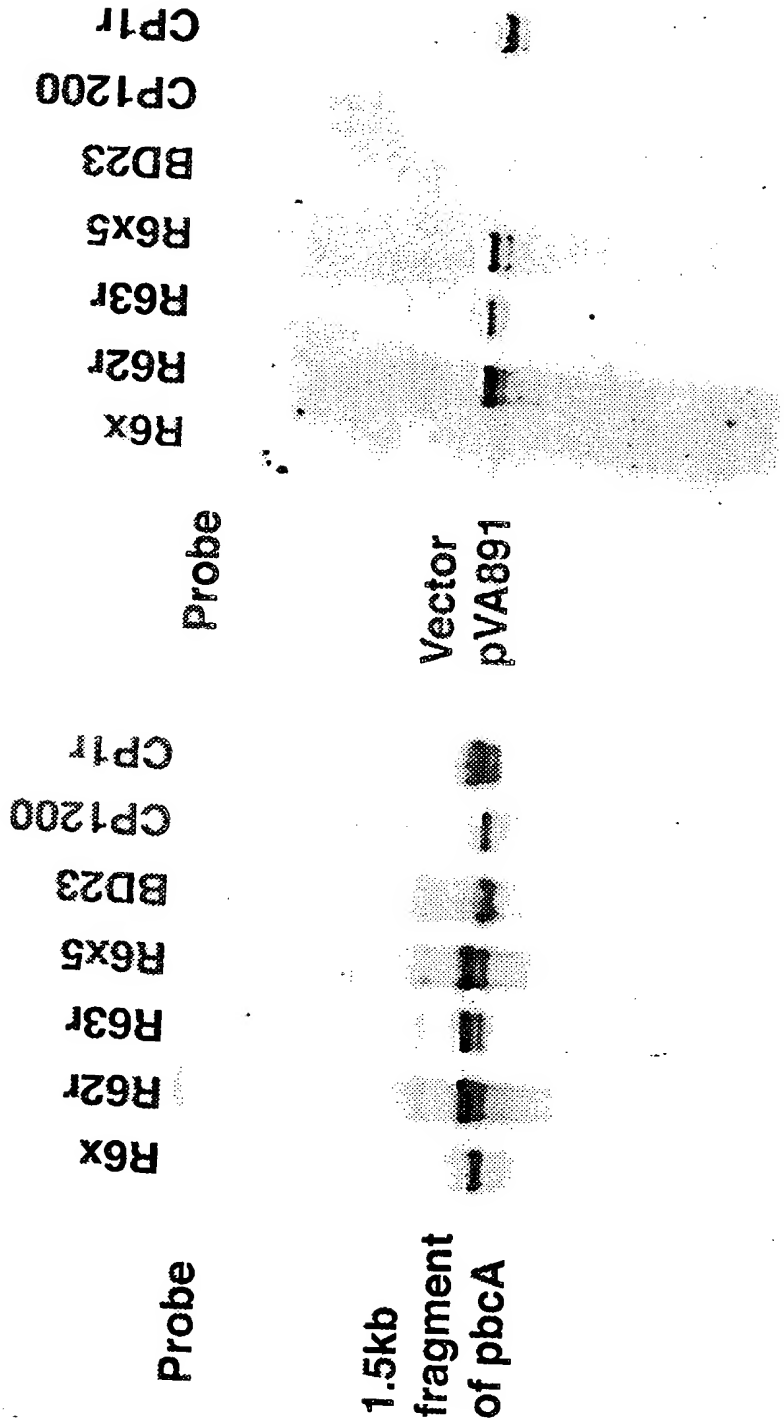
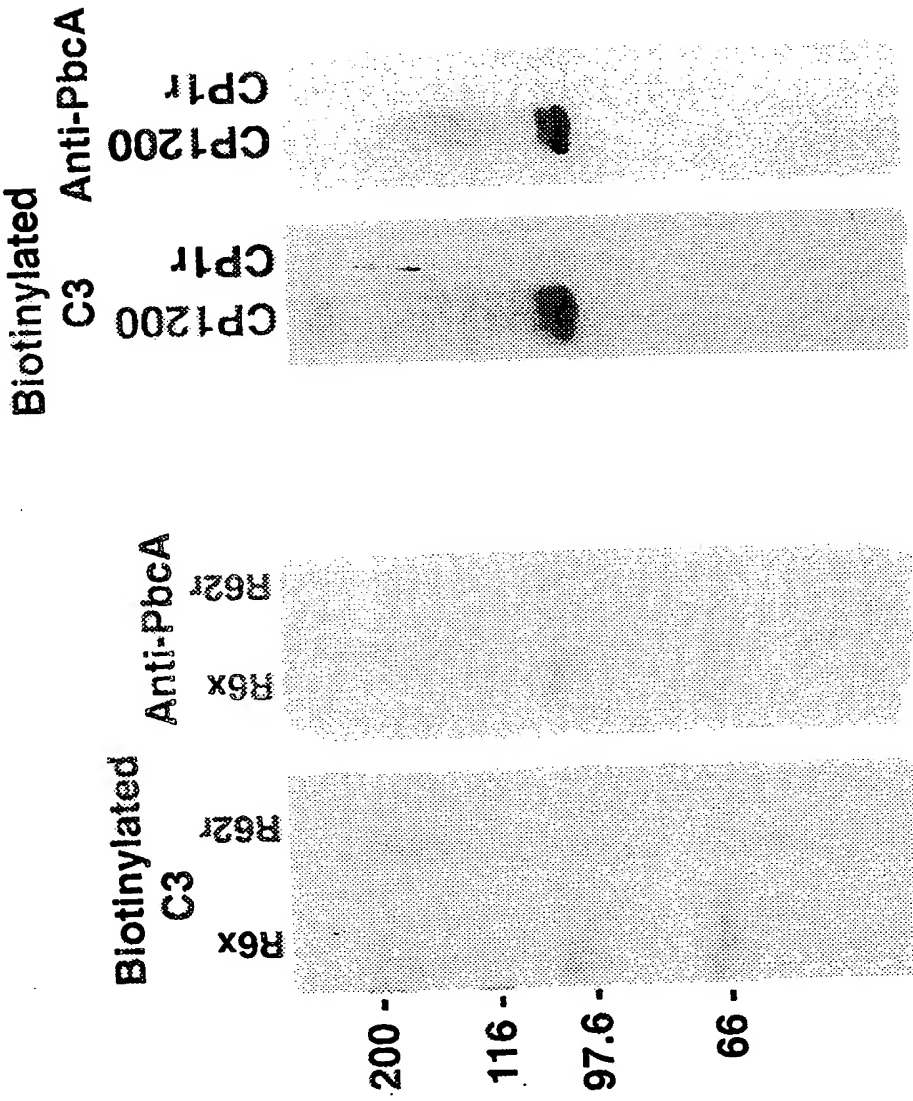


Fig. 5a

Fig. 5b



INTERNATIONAL SEARCH REPORT

Intern. ial Application No
PCT/US 97/20586

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12N15/31 C12N1/21 C07K14/315 C07K16/12 A61K39/09
A61K39/40 //(C12N1/21,C12R1:46)

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12N C07K A61K

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	WO 97 41151 A (UNIV ROCKEFELLER) 6 November 1997 see page 9, line 11 - line 19 see page 67, line 14 - line 25 see page 73, line 18 - page 74, line 8 SEQ ID no.24,25 see page 76, line 15 - page 78, line 13; claims 1-40	1-68
A	EP 0 622 081 A (UAB RESEARCH FOUNDATION) 2 November 1994 see the whole document	1-68
A	WO 93 24000 A (YOTHER JANET) 9 December 1993 see the whole document	1-68

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
8 May 1998	02.06.98

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Hornig, H

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Inter nal Application No
PCT/US 97/20586

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
P,X	ANONYMOUS: "A C3 - binding protein from Streptococcus pneumoniae." 97TH GENERAL MEETING OF THE AMERICAN SOCIETY FOR MICROBIOLOGY, MIAMI BEACH, FLORIDA, USA, MAY 4-8, 1997. ABSTRACTS OF THE GENERAL MEETING OF THE AMERICAN SOCIETY FOR MICROBIOLOGY 97 (0). 1997. 110. ISSN: 1060-2011, XP002064458 Accession no. B-478 ---	1-28, 33-47
P,X	C. ROSENOW ET AL.: "Contribution of novel choline-binding proteins to adherence, colonization and immunogenicity of Streptococcus pneumoniae" EMBL SEQUENCE DATABASE, 29 September 1997, HEIDELBERG, FRG, XP002064459 Accession no. AF019904; AC 030874; ---	1-28, 33-47
P,X	S. HAMMERSCHMIDT ET AL.: "SpsA, a novel pneumococcal surface protein with specific binding to immunglobulin A and secretory componnets" EMBL SEQUENCE DATABSE, 14 October 1997, HEIDELBERG, FRG, XP002064460 Accession no. AJ002054 -----	1-28, 33-47

1

Form PCT/ISA/210 (continuation of second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

Inter ational application No.
PCT/US 97/20586

Box I Observations where certain claims were found unsearchable (Continuation of item 1 of first sheet)

This International Search Report has not been established in respect of certain claims under Article 17(2)(a) for the following reasons:

1. [X] Claims Nos.:
because they relate to subject matter not required to be searched by this Authority, namely:

Although claims 55-63 are directed to a method of treatment of the human/animal body, the search has been carried out and based on the alleged effects of the compound/composition.

2. [ ] Claims Nos.:
because they relate to parts of the International Application that do not comply with the prescribed requirements to such an extent that no meaningful International Search can be carried out, specifically:

3. [ ] Claims Nos.:
because they are dependent claims and are not drafted in accordance with the second and third sentences of Rule 6.4(a).

Box II Observations where unity of invention is lacking (Continuation of item 2 of first sheet)

This International Searching Authority found multiple inventions in this international application, as follows:

1. [ ] As all required additional search fees were timely paid by the applicant, this International Search Report covers all searchable claims.

2. [ ] As all searchable claims could be searched without effort justifying an additional fee, this Authority did not invite payment of any additional fee.

3. [ ] As only some of the required additional search fees were timely paid by the applicant, this International Search Report covers only those claims for which fees were paid, specifically claims Nos.:

4. [ ] No required additional search fees were timely paid by the applicant. Consequently, this International Search Report is restricted to the invention first mentioned in the claims; it is covered by claims Nos.:

Remark on Protest

[ ] The additional search fees were accompanied by the applicant's protest.

[ ] No protest accompanied the payment of additional search fees.

INTERNATIONAL SEARCH REPORT

Information on patent family members

Inter nal Application No
PCT/US 97/20586

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
WO 9741151 A	06-11-1997	AU 2818297 A	19-11-1997
EP 0622081 A	02-11-1994	AU 682018 B	18-09-1997
		AU 5769694 A	27-10-1994
		CA 2116261 A	21-10-1994
		FI 941695 A	21-10-1994
		JP 7126291 A	16-05-1995
		NO 941420 A	21-10-1994
		US 5679768 A	21-10-1997
		ZA 9401584 A	12-10-1994
WO 9324000 A	09-12-1993	AU 4400293 A	30-12-1993